Year : 2016

CORRELATES AND GENETIC BASIS OF PLANT DIVERSIFICATION IN THE NEOTROPICS: THE GESNERIACEAE AS A CASE STUDY

Serrano Serrano Martha Liliana

UNIL | Université de Lausanne
Faculté de biologie et de médecine

Département d'Écologie et d'Évolution

CORRELATES AND GENETIC BASIS OF PLANT DIVERSIFICATION IN THE NEOTROPICS: THE GESNERIACEAE AS A CASE STUDY

Thèse de doctorat ès sciences de la vie (PhD)
Écologie et Évolution

Présenté à la
Faculté de biologie et de médecine de l'Université de Lausanne
par

Martha Liliana SERRANO SERRANO

Master en Biologie et systématique de l'Université Nationale de Colombie
Bogotá, Colombie

Jury

Prof. Christian Fankhauser, président
Prof. Nicolas Salamin, directeur
Prof. Mathieu Perret, co-directeur
Prof. Mark Chapman, expert
Prof. John Pannell, expert

Lausanne 2016

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président · e	Monsieur	Prof. Christian **Fankhauser**
Directeur · rice de thèse	Monsieur	Prof. Nicolas **Salamin**
Co-directeur · rice	Monsieur	Dr Mathieu **Perret**
Experts · es	Monsieur	Prof. John **Pannell**
	Monsieur	Prof. Mark **Chapman**

le Conseil de Faculté autorise l'impression de la thèse de

Madame Martha Liliana Serrano Serrano

Master of Sciences Universidad Nacional de Colombia, Bogota, Colombie

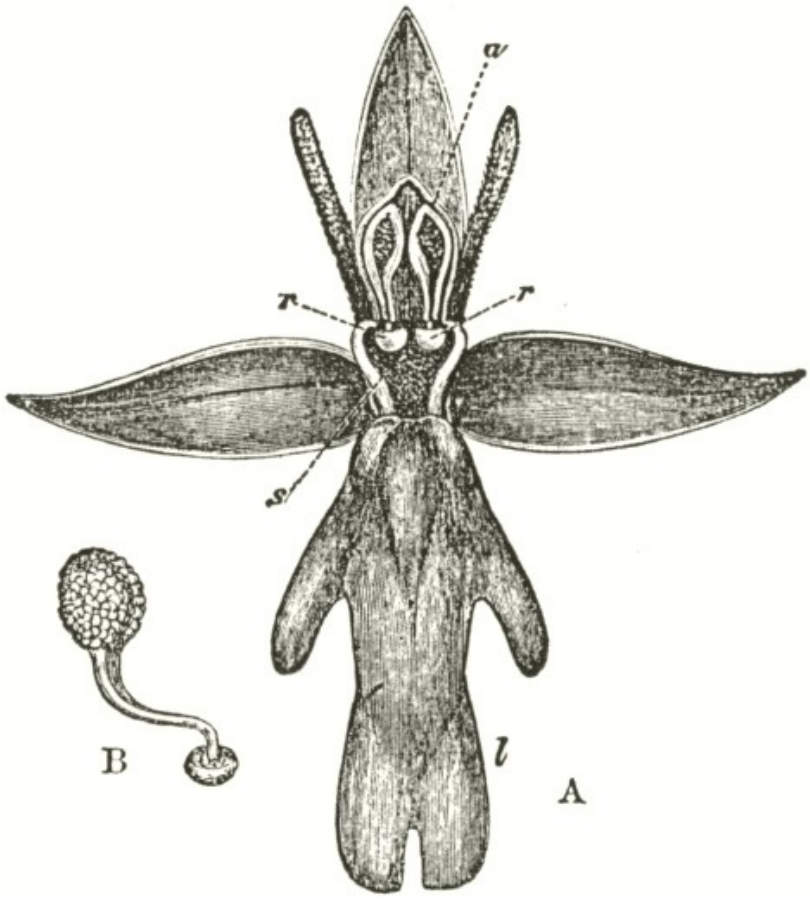
intitulée

CORRELATES AND GENETIC BASIS OF PLANT DIVERSIFICATION IN THE NEOTROPICS : THE GESNERIACEAE AS A CASE STUDY

Lausanne, le 16 septembre 2016

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Christian Fankhauser



Ophrys muscifera (Darwin, 1862)

"Flowers rank amongst the most beautiful productions of nature; but they have been rendered conspicuous on contrast with green leaves, and in consequence at the same time beautiful, so that they may be easily observed by insects"

Darwin 1859.

Table of contents

Abstract

The temporal building of plant biodiversity is a fascinating and extensive research area in evolutionary biology. The striking patterns of species richness in the Neotropical area, and the variety of floral forms, have long challenged our understanding of how plant speciation proceeds, what is required for acquiring reproductive isolation, and which are the factors that may promote it. In this thesis, I combine macro-evolutionary and genetic approaches at an integrative taxonomic and temporal scale to investigate the diversification processes shaping the captivating Gesneriaceae plant family. In the first section, I identify the dynamics of evolution of floral morphologies and climatic preferences within an epiphytic and endemic lineage of the gesneriads. I show a strong decoupling in the evolution of floral shape and size that reflects the contrasting influence of plant-pollinator interactions: while floral shapes converge towards optimal morphologies that are highly bounded to specific bee and hummingbird interactions, floral sizes rapidly diverge once hummingbird pollination is established in a lineage. Building on these findings, I reconstruct a large species-level phylogeny, and evaluate the evolution of plant-pollinator interactions at the scale of the whole subfamily. I demonstrate that transitions between pollinator types are very frequent and symmetrical, suggesting a pattern of evolutionarily labile interactions. Moreover, I identify a positive effect of hummingbird-pollination in the rates of speciation. This pattern contrast with the traditional view on pollination-driven plant speciation, which regards shifts between pollinators as the driver of diversity. In the final section, in order to grasp these patterns at fine genetic scale, I generate the transcriptomic data for six closely related species that underwent pollination shifts. I investigate whether parallel genetic changes have produced the alike floral morphologies associated to the same plant-pollinator interactions. This exploration sheds light on the mechanisms associated with flower transformations in the Gesneriaceae, emphasizing the stronger impact of molecular evolution over gene expression differences, and pinpointing to several candidate genes and functional pathways for future research on the evolution of floral morphology. Overall, this work shows how such a multidisciplinary approach at variable evolutionary scales can contribute to disentangle processes that have generated the enormous Neotropical biodiversity that we witness today.

Résumé

L'émergence de biodiversité chez les plantes est un domaine d'étude fascinant et en plein essor de la biologie évolutive. La remarquable richesse spécifique dans les néo-tropiques, ainsi que la variété des formes florales ont longtemps posé un défi pour la compréhension des processus de spéciation chez les plantes, en particulier l'établissement de l'isolement reproductif et les facteurs qui le promeuvent. Dans cette thèse, je combine des approches macro-évolutives et génétiques à différentes échelles taxonomiques et temporelles afin d'étudier les processus de diversifications ayant lieu dans la captivante famille des Gesnériacées. Dans la première section, j'identifie la dynamique évolutive des morphologies florales et des préférences climatiques au sein d'une lignée épiphyte. Je montre un fort découplage entre l'évolution de la taille et la forme des fleurs, qui reflète l'influence contrastée des interactions plantes-pollinisateur: alors que les formes florales convergent vers des morphologies optimales fortement liées à des interactions spécifiques avec des abeilles et des colibris, la taille des fleurs diverge rapidement une fois que la pollinisation par un colibri est établie dans une lignée. Tirant parti de ces résultats, je reconstruis une large phylogénie de la famille et évalue l'évolution des interactions plantes-pollinisateurs à l'échelle de l'ensemble de la sous-famille. Je démontre que les transitions entre types de pollinisateurs sont très fréquentes et symétriques, ce qui suggère des interactions évolutives labiles. De plus, j'identifie un effet positif de la pollinisation par les colibris sur le taux de spéciation. Ces résultats contrastent avec la vue traditionnelle qui considère les transitions entre pollinisateurs comme générateur de diversité. Dans la section finale, afin d'identifier ces patrons à l'échelle génétique fine, je génère des données transcriptomiques pour six espèces apparentées ayant subi des transitions de pollinisateurs. J'étudie si des changements génétiques parallèles ont produits des morphologies florales similaires au sein des mêmes interactions plante-pollinisateur. Cette exploration nous informe sur les mécanismes associés avec la transformation des morphologies florales au sein des Gesnériacées. Elle souligne notamment la prépondérance de l'impact de l'évolution moléculaire sur les différences d'expression de gènes et met en lumière plusieurs gènes candidats et voies fonctionnelles pour de futures recherches sur l'évolution des morphologies florales. En général, ce travail illustre comment une telle approche multidisciplinaire à diverses échelles évolutives peut contribuer à désenchevêtrer les processus qui ont généré l'énorme biodiversité néo-tropicale dont nous sommes témoin aujourd'hui.

General introduction

Insights into a spectacular natural phenomena: plant species diversification

Angiosperms are a remarkable case of rich group encompassing tremendous biodiversity that has long attracted the attention of scientists. One of the key questions that has always puzzled biologists since the early days of evolutionary biology is why are the angiosperms so species-rich and ecologically successful (Darwin, 1871). Darwin's descriptions of angiosperm diversity pointed towards two remarkable directions: the "abominable mystery" and the "beautiful contrivances" (Darwin, 1859). The "abominable mystery" may symbolize the extraordinary rapidity with which flowering plants seem to have evolved. The dynamic of species appearance in angiosperms suggests instances of apparent speed-ups in the rate of speciation, which led to speculations about an effect of co-evolutionary interactions between plants and pollinators, in particular insects (Friedman, 2009). Flowering plants have also been described by Darwin as "beautiful contrivances" and he reached this conclusion after large sessions of observations and experimentation on orchid flowers, leading to the recognition of a remarkable phenotypic diversity, with a potential adaptive nature (Harder & Johnson, 2009). These two ideas suggest that present day diversity of angiosperms is the product of a complex evolutionary process, in which lineages split favored by ecological interactions, such as the pollination function, while diversifying in a broad mosaic of forms. This premise drive us to some of the most fascinating evolutionary research areas: the inferences of tempo and mode of species diversification, the pollinators as drivers of plant diversification, and the evolution and development of floral diversity.

The integration of these three areas of research provides the theoretical and methodological background for advancing our understanding of plant speciation. The first area introduces the study of the tempo and mode of evolution in a lineage, which combines phylogenetic and trait comparative methods, to statistically infer the macroevolutionary dynamics in a group of related organisms (Felsenstein, 1988; Harvey & Pagel, 1991). These methods help to elucidate the speed, timing, and direction of evolutionary events, such as the origination and rate of evolution of new traits, the rates of species diversification, and the association between trait evolution and their effects on speciation and extinction (Maddison *et al.*, 2007; Freckleton *et al.*, 2008). The second area drives the attention on how ecological factors, specifically ecological interactions such as plant-pollinator relationships, may promote speciation in angiosperms (Vamosi & Vamosi, 2011). One widely supported beneficial factor is the relationship between plants and their pollinators (Crepet & Niklas, 2009; van der Niet & Johnson, 2012). The mechanisms by which plant-pollinator interactions increase the probability of speciation include the opportunities for floral diversification, and the impact of this variation on reproductive isolation (Kay & Sargent, 2009; Baack *et al.*, 2015). Consequently, the third area complements the emerging picture of angiosperm evolution by addressing the importance of floral evolution, and the ability of a lineage to diversify morphologically (Smith, 2010; Armbruster, 2014). Recent advances pinpointed that the genetic basis of floral changes may determine the ability and extent of phenotypic variation in a lineage (Preston *et al.*, 2011). The simultaneous investigation of these three aspects in a species-rich family of angiosperms remains unexplored, and it represents an excellent opportunity to merge distinct evolutionary processes within a multidisciplinary framework. The ability to explore these three axes of angiosperm evolution can contribute to answer a longstanding enigma that puzzled Darwin, the processes underlying the rich and enormous plant diversity in the Neotropics:

In England any person fond of natural history enjoys in his walks a great advantage, by always having something to attract his attention; but in these fertile lands teeming with life, the attractions are so numerous, that he is scarcely able to walk at all.

Charles R. Darwin, 19 Apr. 1839, after leaving Brazil on board of HMS Beagle

Inferences of tempo and mode of species diversification

Understanding what shapes biodiversity and large-scale patterns of species richness are some of the most challenging research questions in the fields of evolution, ecology, genetics and paleontology (Pennisi, 2005). Consequently, the study of species formation is a key aspect to identify the factors that promote biodiversity. The study of speciation is favored, in a simplified way, by the examination of the levels of reproductive isolation, morphological distinctness, and phylogenetic relationships between groups (Coyne & Orr, 2004). In practice, the study of speciation at the macroevolutionary scale has been motivated by the phylogenetic evidence. A species-level phylogeny is an inference of the relationships between species, and provides information about the underlying branching process through time (Hillis *et al.*, 1996). This information can be transformed into estimates of diversification rates by using birth-death stochastic processes, which model the effects of speciation and extinction within a lineage (Nee, 2001; Nee, 2006; Stadler, 2013). The heterogeneity of these two processes across lineages leads to the first striking pattern in biodiversity: the noticeable disparity in species richness that is observed in different groups of organisms (Ricklefs, 2007). A difference in the number of species between similar-aged lineages is most probably the product of heterogeneity in diversification rates (Figure 1a), which are due to a variety of factors (e.g. traits) that may affect either speciation and/or extinction processes.

Relating species traits to the differences in diversification process is one of the main aims of comparative methods (Freckleton *et al.*, 2008). Intuitively, a first step on the exploration of these differences is to propose a hypothesis relating a trait, and its biological mechanism, with its effects on rates of diversification. In angiosperms, a variety of traits have been proposed to affect the speciation and extinction processes (Antonelli & Sanmartín, 2011; Vamosi & Vamosi, 2011). They include nectar spurs (Hodges & Arnold, 1995; Hodges, 1997), floral characters and specialization (Sargent, 2004; Kay *et al.*, 2006; Armbruster & Muchhala, 2009), defense traits (Agrawal *et al.*, 2009), type of fruits (Smith, 2001), pollination syndromes (Dodd *et al.*, 1999; Kay *et al.*, 2005), sexual systems (Goldberg & Igic, 2012), and growth habit (Gianoli, 2004; Gravendeel *et al.*, 2004). Using the trait and phylogenetic information, one can statistically infer their historical patterns to discover how they have evolved, the timing of origination, and even the correlated evolution with any additional trait (Pagel, 1994; Huelsenbeck *et al.*, 2003). However, the independent study of the evolution of a trait, even in a phylogenetic context, does not guarantee the proper association with the diversification process, and overall does not tease apart asymmetries in diversification from asymmetries in trait change (Maddison, 2006; Maddison & FitzJohn, 2015). Over the last ten years, methodological advances have allowed us to jointly estimate these two evolutionary processes, to explain the observed distribution of traits and the richness patterns (O'Meara & Beaulieu, 2016).

Maddison *et al.* (2007) first proposed the basis for a family of models, collectively called state-dependent models of diversification, that jointly estimate the rates of transitions for a binary trait and diversification. The initial model allows the integration of all possible scenarios along a branch in a phylogenetic tree (i.e. trait changes or stasis, speciation, extinction, or none), and based on the probabilities of the observed tree and trait states, estimates six evolutionary parameters (rates of diversification and state transition, see Figure 1b). Later extensions of this model included the evaluation with multi-state and continuous traits (FitzJohn, 2010; FitzJohn, 2012), geographic scenarios (Goldberg *et al.*, 2011), and the distinction of trait changes associated to speciation events or along single branches (Goldberg & Igic, 2012). Multiple criticisms and potential biases have been identified for this family of models (Davis *et al.*, 2013; Maddison & FitzJohn, 2015), but the examination of traits that have changed multiples times within a lineage remains a robust way to test for trait effects in diversification (O'Meara & Beaulieu, 2016).

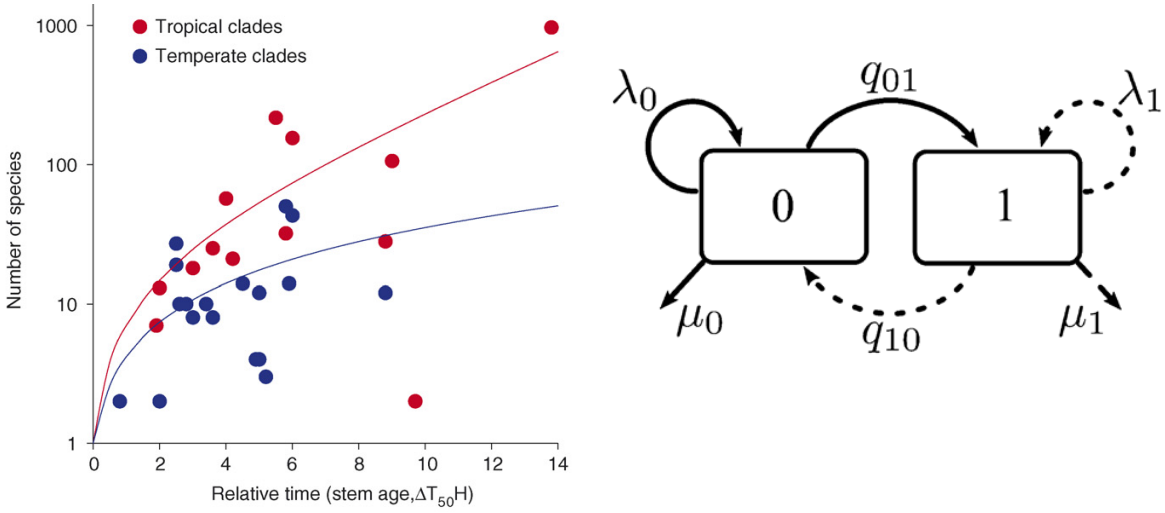


Figure 1. A, Maximum likelihood estimates of the relationship between number of species and relative age for South American clades (tropical, red dots) and North American clades (temperate, blue dots) of passerine birds (Ricklefs, 2007). B, The states and allowed transitions in BiSSE models. Six rate parameters represent the state-dependent speciation (λ) and extinction (μ), and the state transition rates (image taken from Goldberg *et al.*, 2011).

Pollinators as drivers of plant diversification

Plant and pollinator relationship is one of the most ecologically important plant-animal interactions. This mutualistic relationship dates back to the early stages of the angiosperms evolution (Michez *et al.*, 2012), and seems to be highly successful with up to a 94% of plant species pollinated by animals in tropical communities (Ollerton *et al.*, 2011). The association with animal pollinators seems to have driven the rapid diversification of flowering plants, with a potential effect on the opportunities for floral specialization and diversification (Dodd *et al.*, 1999; Crepet & Niklas, 2009). However, what is the role of pollinators in speciation? Speciation requires the acquisition of reproductive isolation among populations, and the divergences due to pollination interactions could promote this partitioning and isolation in multiple ways.

Reproductive isolation driven by pollinator interactions is long been suggested to be the result of floral isolation (Grant, 1949). The mechanisms for acquiring this floral isolation can include varied ecological

scenarios. First, imagine a population distributed over a geographical range that presents spatial differences in pollinator abundances. While one type of pollinator is more frequent in a site, a different one is dominating the second part of the species range. This long term interaction could eventually drive the evolution of distinct floral traits, increasing the specificity for each pollinator, and reducing the level of gene flow until reaching speciation (Grant & Grant, 1965). The persistence of this scenario over evolutionary time could be the underlying cause of speciation in a plant lineage (Whittall & Hodges, 2007; Forest *et al.*, 2014; Van der Niet *et al.*, 2014; Breitkopf *et al.*, 2015). Alternatively, floral isolation could be driven by pollinator competition in sympatry. Given that plant species aim to efficiently transfer pollen to their conspecifics, the use of the same pollen vector (i.e. functional group of pollinators) may drive floral character displacement and mechanical isolation, or timing and behavioral differences, to reduce interspecific pollination (Armbruster *et al.*, 1994; Muchhala & Potts, 2007; Muchhala *et al.*, 2014). Overall, there are more studies investigating the changes in pollination systems and floral traits, than cases of floral divergences within a pollinator type, and for the later the contribution on species richness has received less attention.

Recent literature highlighted the value of investigating the pollinator-driven speciation process at different evolutionary scales and wide taxonomic range (Special issue of Annals of Botany journal in 2014: Volume 113 Issue 2). Sources of evidence for pollinator-driven diversification and reproductive isolation have been more extensively explored at the population and closely related species levels (Ramsey *et al.*, 2003; Schiestl & Schluter, 2009; Hopkins *et al.*, 2011). Taking a historical perspective that combines phylogenetic and comparative methods would help to get a better picture of the frequency and direction of pollinator shifts within a lineage, the level of floral divergence between species, and the precise test for a contribution to species richness (van der Niet & Johnson, 2012). Altogether these components offer a promising framework to answer diverse questions on the role of pollinators in angiosperm diversification (Table 1).

Table 1. Central questions regarding the role of pollinators in angiosperm diversification (modified from Van der Niet & Johnson, 2012). The questions that have received a contribution from this thesis are marked in the last column.

Topic	Question	Approach	Results
Pollinators and angiosperm diversification	Are pollinator shifts correlated with floral shifts?	Testing whether pollinator shifts and floral shifts evolve independently	X
		Using correlations between pollinator number or contribution with floral traits, while accounting for phylogeny	
	What is the frequency of convergent pollinator shifts?	Optimization or mapping of pollinators onto a phylogeny and counting independent origins	X
		Studying homology of morphological characters associated with convergent origins of pollination systems	
	Are pollinator shifts associated with shifts in plant diversification rates?	Using an evolutionary model that estimates transition rates between pollination systems and plant speciation and extinction rates	X
		Assessing whether branches with a significant change in plant diversification rates are characterized by pollinator shifts	
The factors governing the likelihood of pollinator shifts	Are pollinator shifts correlated with shifts in plant breeding systems?	Testing whether the level of pollinator specialization is correlated with the level of autofertility	
	What is the importance of pollinator shifts relative to other ecological shifts that promote plant diversification?	Comparing pollinators and ecological attributes of sister species	
	What is the direction of pollinator shifts?	Optimization or mapping of pollinators onto a phylogeny and assessing the shift direction	X
		Using an evolutionary model to estimate transition rates between pollination systems	X
	Are pollinator shifts determined by changes in the local pollinator assemblage?	Comparing the ranges of pollinators and plants for plant sister species pairs	
Pollinators and plant speciation	Are pollinator shifts associated with plant speciation?	Testing whether pollinator shifts occurred during or after speciation	
	Do pollinator shifts evolve as a means of reproductive isolation in sympatric plant species?	Comparing isolating mechanisms between plant sister species in relation to distribution ranges	

Evolution and development of floral diversity

Flower morphologies are hypothesized to be the result of similar selective forces exerted by pollinators, and they may reflect the level of specialization onto specific groups of pollinators. This specialization has been experimentally tested by evaluating pollination preferences in an array of floral traits, suggesting a clear effect of floral traits in pollinator visitation rates (Figure 2). The combination of traits promoting a specific plant-pollinator interaction is called "pollination syndrome" (Fenster *et al.*, 2004). However, the stability and accuracy for the prediction of pollinators from pollination syndrome morphologies are highly debated (Ollerton *et al.*, 2009). The reason is because floral evolution is not only increasing the specialization towards a single pollinator, but rather an array of those (Ashworth *et al.*, 2015), while simultaneously avoiding other visitors (Castellanos *et al.*, 2004; Cronk & Ojeda, 2008). Additionally, multiple antagonistic forces, such as stress tolerance and plant defenses, may impose selection on floral variability (Strauss & Whittall, 2006). Despite these conflicting forces, the study of transitions between pollinators has shown that flower morphologies are highly convergent, as if phenotypes were directed towards optimal adaptive peaks, and jumps to new pollinator interactions usually involve coordinated multivariate character changes (Thomson & Wilson, 2008). Phylogenetic evidence of these shifts has suggested that considerable floral trait changes are required on the process (e.g. in flower color and nectar rewards), and those shifts are potentially affecting the speciation dynamic of a lineage (Wilson *et al.*, 2006; Smith, 2010). Consequently, to understand the evolutionary value of floral diversity, it is crucial to identify the genetic basis for floral changes, the number and location of loci involved, their structural or regulatory nature, and the magnitude of their phenotypic effects (Kay & Sargent, 2009).

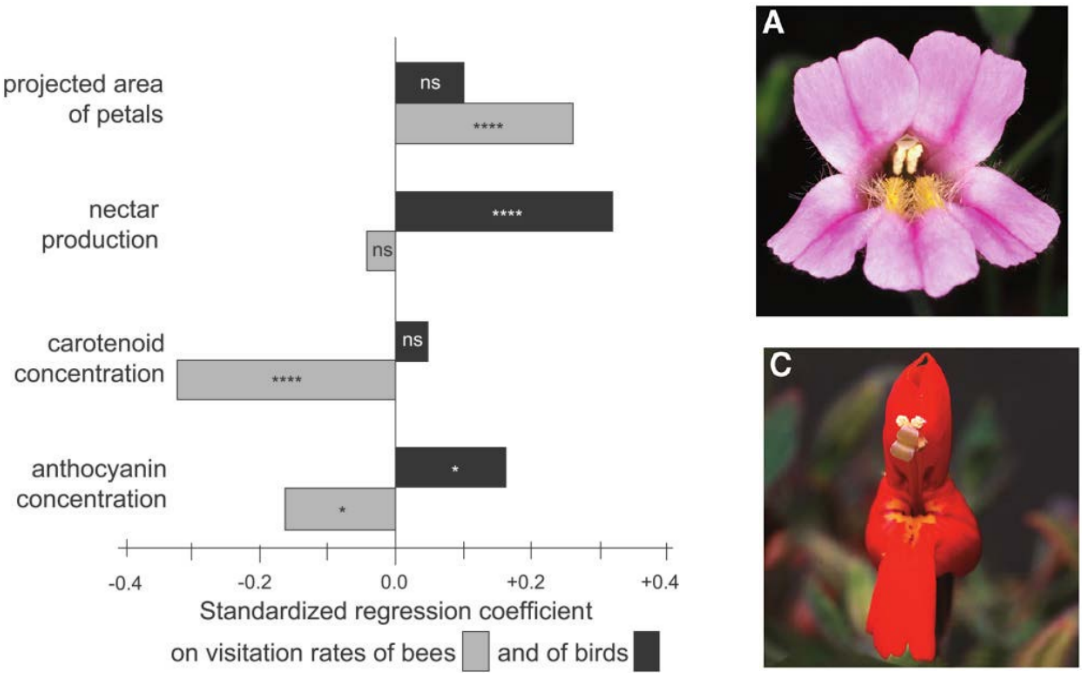


Figure 2. Multiple regression analysis of four floral traits on visitation rates by bees and birds in an F2 array of monkeyflowers (A, *Mimulus lewisii*, and B, *M. cardinalis*). Figure from Thomson and Wilson (2008), and pictures from Schemske and Bradshaw (1999).

Recent studies have specifically targeted the genes controlling relevant pollinator-attraction floral traits, named "pollination syndrome genes" (Stuurman *et al.*, 2004; Galliot *et al.*, 2006; Hermann *et al.*, 2015). The genetic study of isolated floral traits, such as flower color, has received considerable attention over the last decade resulting in a detailed understanding of its genetic basis in multiple model species (Bradshaw & Schemske, 2003; Whittall *et al.*, 2006; Smith & Rausher, 2011; Wessinger & Rausher, 2014). However, multiple morphological changes including petal color, scent, corolla shape, and nectar production among others, seem to happen coordinately during pollination shifts (Thomson & Wilson, 2008). Studies conducted in species from the genus *Petunia* (Solanaceae) have shown that a relatively simple genetic control, with few major QTLs (Quantitative Trait Loci) or tight genetic linkage regions, underlie floral scent and anthocyanin regulation, UV absorbance, and pistil and stamen length (Klahre *et al.*, 2011; Hermann *et al.*, 2013; Sheehan *et al.*, 2016). Most of these studies have been conducted on model systems (see Wessinger *et al.*, 2014 for findings in species from Penstemon genus), and the identification of the genetic control in an evolutionary context is in an early stage (Cronk & Ojeda, 2008).

Progress towards an understanding of the coordinated evolution of flowers and pollinators, and the repeatability of this process has been done by combining molecular and phylogenetic frameworks. The examination of the genetic changes during pollination transitions in the genus *Ipomoea* (Convolvulaceae) has revealed that shared changes in gene expression, and molecular evolution of specific genes, have shaped the evolution of similar morphologies within this lineage (Des Marais & Rausher, 2010; Smith *et al.*, 2013). Moreover, the advances in whole-genome, transcriptome and proteome sequencing techniques are promising in the identification of genes related with floral morphologies (Lulin *et al.*, 2012; Sedeek *et al.*, 2013; Singh *et al.*, 2013). These approaches provide an excellent opportunity to explore candidate genes, especially in non-model organisms, where no extensive genetic resources are available, but with a relevant value for ecological and evolutionary research (Elmer & Meyer, 2011).

The Gesneriaceae study system

Testing the contribution of pollinators and floral evolution into the building of plant biodiversity is facilitated by the identification of a study group that provides sufficient morphological variability, and intriguing ecological and biodiversity patterns. Most of these biological groups are natural populations and non-model species, which remain largely unexplored (Savolainen *et al.*, 2013). However, the rapid development of genomic tools for accessing to the molecular information in these species (Todd *et al.*, 2016), and the thoughtful integrative approaches for understanding the patterns of plant biodiversity (Antonelli & Sanmartín, 2011; Givnish *et al.*, 2014), have facilitated the hypothesis testing in rather unexplored taxonomic groups. Here, I combined phylogenetic information, comparative trait evolution models, and cutting-edge sequencing technologies for advancing our understanding of the speed and mode of character evolution, their impact on the speciation process, and the genetic mechanisms associated to the evolution of the Gesneriaceae family.

The neotropical lineage of the Gesneriaceae family, classified as the Gesnerioideae subfamily, contains around 1000 species, divided into five tribes and 12 subtribes (Weber *et al.*, 2013). The subfamily is a promising study system because of its wide distribution over the Neotropical region (Figure 3), and the high morphological diversity. It represents an excellent opportunity to merge different types of data and disciplines to explore the factors affecting biodiversity. Traditional characters used to define and classify taxa are highly labile in Gesnerioideae, though current phylogenetic evidence derived from molecular data has resolved monophyletic lineages in the group (Möller & Clark, 2013), making possible sound

taxonomic revisions for many genera and tribes (Roalson *et al.*, 2008; Chautems *et al.*, 2010; Clark *et al.*, 2012; Perret *et al.*, 2013; Smith & Clark, 2013; Araujo *et al.*, 2016; Ferreira *et al.*, 2016; Mora & Clark, 2016). Taxonomic reviews evidenced a large disparity in species richness across clades, with a few genera including several hundred species and a larger number of genera are only represented by a single or few species (Weber *et al.* 2013). The evolution of traits such as fruit morphologies, pollination systems, growth habit types, as well as the distinct geographical distribution, have been investigated in the Gesnerioideae (Smith, 2000; Perret *et al.*, 2003; Martén-Rodríguez *et al.*, 2009; Clark *et al.*, 2012), but their contribution to the richness imbalance has been narrowly tested (Smith, 2001).

The species in the family Gesneriaceae have a notorious floral diversity and a wide range of interactions with pollinators (Figure 4). These interactions include insects (mostly bees), hummingbirds and bats (Table S2 in Chapter 2). The morphological characterization of the subtribes Ligeriinae and Gesneriieae have identified strong correlations between floral phenotypes and functional groups of pollinators (Perret *et al.*, 2001; Perret *et al.*, 2007; Martén-Rodríguez *et al.*, 2009; Chautems *et al.*, 2010), in many cases validated by observations in the field. Consequently, floral morphologies may represent reliable predictors of pollination systems, even when a system rather generalized than specialized in a single pollinator (Fenster *et al.*, 2004). Studies on several Gesnerioideae lineages have suggested a labile pattern of plant-pollinator associations, with the presence of multiple pollinator types within a single lineage (Perret *et al.*, 2003; Marten-Rodriguez *et al.*, 2010), but an analysis at a large taxonomic scale is lacking. The presence of independent pollinator shifts during the history of the Gesnerioideae subfamily provides a replicated evolutionary framework to investigate the speed of those changes, and the parallel nature of the genetic mechanisms (Elmer & Meyer, 2011; Martin & Orgogozo, 2013).

The objectives of this thesis

At a broad scale, the main goal of this thesis was to investigate the diversification patterns of the Gesneriaceae family in the Neotropics, by first evaluating within a phylogenetic comparative framework the impact of pollination types on morphological evolution and species diversification, and second by examining the molecular signatures (gene expression and sequence differences) during replicated transitions between bee and hummingbird pollination interactions. More specifically:

In **chapter 1,** I investigated the dynamics of trait evolution for floral morphology and climatic preferences, within an epiphytic lineage that includes the *Codonanthopsis-Codonanthe-Nematanthus* genera. I provided evidence of a decoupled evolutionary dynamic between all the traits examined. The evolution of flower shape reflected a strong association with pollination types in the group, with a very constrained diversification pattern. In contrast, flower size underwent a rapid evolution in a specific clade associated with the transition to hummingbird pollination. Climatic preferences in the lineage showed a separation between neotropical biomes, with very few climatic variables deviating from a constant pace of evolution.

In **chapter 2,** I examined whether plant-pollinator interactions have played a role in the species diversification process at the Gesnerioideae subfamily level. I reconstructed a very well sampled species-level phylogeny, to estimate the transition rates between bee and hummingbird pollinator types, and test an effect of each pollinator type on the rates of speciation and extinction. I found that pollination systems are highly labile during the evolutionary history of the subfamily, and that hummingbird pollination is associated with a significant increase in the rates of speciation.

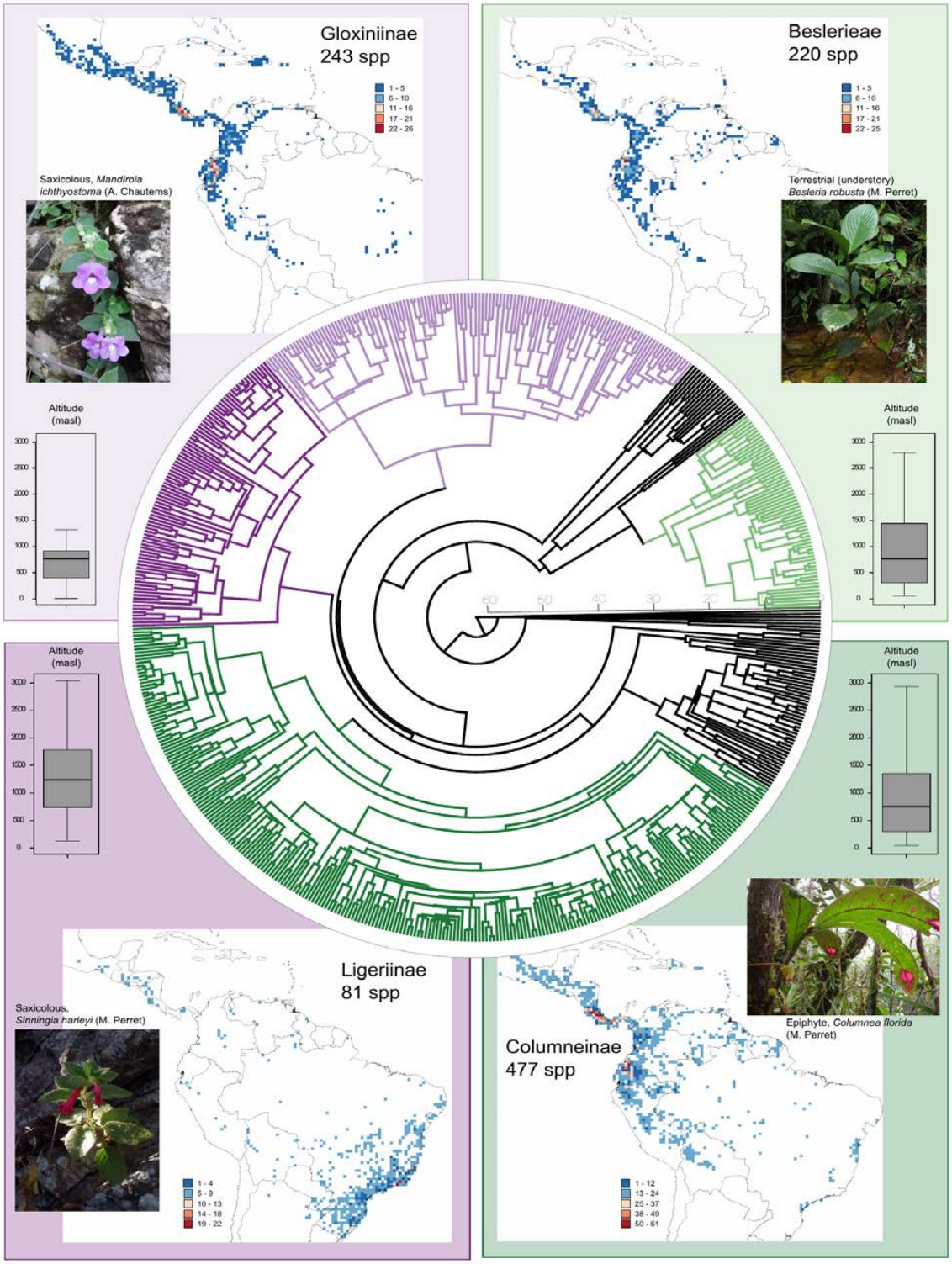


Figure 3. Bayesian phylogeny of the Gesnerioideae subfamily (reconstruction and dating details see chapter 2), and species distribution for the main subtribes (GBIF data). Color scheme represent the species richness according to each inset scale. Pictures provide examples of the growth habit type (species names in italics and picture acknowledgments in brackets).

In **chapter 3,** I built the transcriptomic resources for six related non-model species within the *Nematanthus* and *Sinningia* genera. I gathered next-generation sequencing data aiming at the characterization of three replicated transitions between bee and hummingbird pollination, within a floral developmental framework. I performed the assembly of reference transcriptomes, protein prediction, and annotation for each species. All large scale information was prepared for downstream analyses, and incorporated into a relational database.

In **chapter 4,** I investigated the genetic basis of three replicated pollination shifts in Gesneriaceae family. In particular, I assessed whether gene expression differences and signatures of selection differentially contributed to the establishment of repeated floral morphologies. I found a stronger impact of the signatures of selection, than gene expression differences, in the mechanisms contributing to the genetic control of the concerted floral morphologies in the family. I additionally provided a reduced list of candidate genes for further experimental and evolutionary explorations.

The two annexes of this thesis contain collaborative manuscripts on which I had the opportunity to contribute. They tackle ecological and evolutionary questions, but they are not directly related to the chapters presented above.

Annex 1: "The simultaneous inducibility of phytochemicals related to plant direct and indirect defences against herbivores is stronger at low elevation". Pellissier *et al.* (2016). This paper explores the effects of elevation on the expression of plant defenses in a group of 16 *Cardamine* species, while incorporating the phylogenetic relatedness and the evolutionary history of the group.

Annex 2: "Biogeography and diversification of the New World thatch palms (Cryosophileae and Sabaleae: Arecaceae)". Cano *et al.* in prep. This paper investigates the biogeographic and diversification history of two tribes of palms, with contrasting richness patterns, in the Americas.

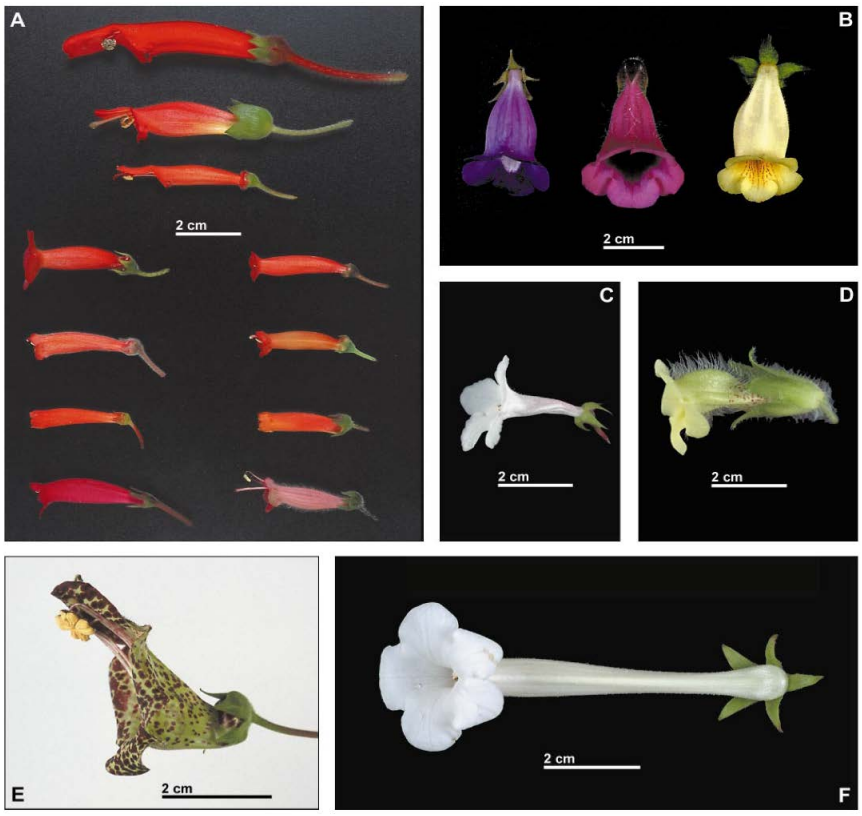


Figure 4. Examples of floral variation within the *Sinningia* genus (from Perret *et al.*, 2001). A, hummingbird flowers (bottom, left and right rows: S. lineata, *S. leucotricha*, *S. insularis*, *S. reitzii*, *S. macropoda*, *S. macrostachya*, *S. aggregata*, and *S. selowii*. Species from top to bottom *S. cardinalis*, *S. incarnata*, *S. bulbosa*); B, C, and D, bee flowers (B: from left to right *S. speciosa*, *S. aghensis*, *S. conspicua*. C: *S. schiineri*, D: *S. villosa*; E, bat flower (*S. brasiliensis*); F, moth flower (*S. tubiflora*).

References

Agrawal AA, Fishbein M, Halitschke R, Hastings AP, Rabosky DL, Rasmann S. 2009. Evidence for adaptive radiation from a phylogenetic study of plant defenses. *Proc Natl Acad Sci U S A* **106**(43): 18067-18072.

Antonelli A, Sanmartín I. 2011. Why are there so many plant species in the Neotropics? *Taxon* **60**(2): 403-414.

Araujo AOd, Chautems A, Cardoso-Gustavson P, Souza VC, Perret M. 2016. Taxonomic Revision and Phylogenetic Position of the Brazilian Endemic Genus Sphaerorrhiza (Sphaerorrhizinae, Gesneriaceae) Including Two New Species. *Systematic Botany*.

Armbruster WS. 2014. Floral specialization and angiosperm diversity: phenotypic divergence, fitness trade-offs and realized pollination accuracy. *AoB Plants* **6**(0).

Armbruster WS, Edwards ME, Debevec EM. 1994. Floral character displacement generates assemblage structure of Western Australian triggerplants (Stylidium). *Ecology* **75**(2): 315-329.

Armbruster WS, Muchhala N. 2009. Associations between floral specialization and species diversity: cause, effect, or correlation? *Evolutionary Ecology* **23**(1): 159-179.

Ashworth L, Aguilar R, Martén-Rodríguez S, Lopezaraiza-Mikel M, Avila-Sakar G, Rosas-Guerrero V, Quesada M 2015. Pollination syndromes: a global pattern of convergent evolution driven by the most effective pollinator. *Evolutionary Biology: Biodiversification from Genotype to Phenotype*: Springer, 203-224.

Baack E, Melo MC, Rieseberg LH, Ortiz-Barrientos D. 2015. The origins of reproductive isolation in plants. *New Phytol* **207**(4): 968-984.

Bradshaw H, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**(6963): 176-178.

Breitkopf H, Onstein RE, Cafasso D, Schluter PM, Cozzolino S. 2015. Multiple shifts to different pollinators fuelled rapid diversification in sexually deceptive Ophrys orchids. *New Phytol* **207**(2): 377-389.

Castellanos MC, Wilson P, Thomson JD. 2004. 'Anti-bee' and 'pro-bird' changes during the evolution of hummingbird pollination in *Penstemon* flowers. *J Evol Biol* **17**(4): 876-885.

Chautems A, Lopes TCC, Peixoto M, Rossini J. 2010. Taxonomic Revision of Sinningia Nees (Gesneriaceae) IV: Six New Species from Brazil and a Long Overlooked Taxon. *Candollea* **65**(2): 241-266.

Clark JL, Funke MM, Duffy AM, Smith JF. 2012. Phylogeny of a Neotropical Clade in the Gesneriaceae: More Tales of Convergent Evolution. *International Journal of Plant Sciences* **173**(8): 894-916.

Coyne J, Orr H. 2004. Speciation. Sinauer. *Sunderland, MA*.

Crepet WL, Niklas KJ. 2009. Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany* **96**(1): 366-381.

Cronk Q, Ojeda I. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *J Exp Bot* **59**(4): 715-727.

Darwin C. 1859. On the origin of the species by natural selection.

Darwin C. 1871. The descent of man. *The Great Books ofthe Western World* **49**: 320.

Davis MP, Midford PE, Maddison W. 2013. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology* **13**(1): 1.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in Ipomoea. *Evolution* **64**(7): 2044-2054.

Dodd ME, Silvertown J, Chase MW. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution*: 732-744.

Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol Evol* **26**(6): 298-306.

Felsenstein J. 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*: 445-471.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination Syndromes and Floral Specialization. *Annual Review of Ecology, Evolution, and Systematics* **35**(1): 375-403.

Ferreira GE, Chautems A, Hopkins MJ, Perret M. 2016. Independent evolution of pouched flowers in the Amazon is supported by the discovery of a new species of Lesia (Gesneriaceae) from Serra do Aracá tepui in Brazil. *Plant Systematics and Evolution*: 1-11.

FitzJohn RG. 2010. Quantitative traits and diversification. *Syst Biol* **59**(6): 619-633.

FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**(6): 1084-1092.

Forest F, Goldblatt P, Manning JC, Baker D, Colville JF, Devey DS, Jose S, Kaye M, Buerki S. 2014. Pollinator shifts as triggers of speciation in painted petal irises (Lapeirousia: Iridaceae). *Ann Bot* **113**(2): 357-371.

Freckleton RP, Phillimore AB, Pagel M. 2008. Relating traits to diversification: a simple test. *Am Nat* **172**(1): 102-115.

Friedman WE. 2009. The meaning of Darwin's 'abominable mystery'. *Am J Bot* **96**(1): 5-21.

Galliot C, Stuurman J, Kuhlemeier C. 2006. The genetic dissection of floral pollination syndromes. *Curr Opin Plant Biol* **9**(1): 78-82.

Gianoli E. 2004. Evolution of a climbing habit promotes diversification in flowering plants. *Proc Biol Sci* **271**(1552): 2011-2015.

Givnish TJ, Barfuss MH, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JA, et al. 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Mol Phylogenet Evol* **71**: 55-78.

Goldberg EE, Igic B. 2012. Tempo and mode in plant breeding system evolution. *Evolution* **66**(12): 3701-3709.

Goldberg EE, Lancaster LT, Ree RH. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst Biol* **60**(4): 451-465.

Grant V. 1949. Pollination systems as isolating mechanisms in angiosperms. *Evolution*: 82-97.

Grant V, Grant KA. 1965. Flower pollination in the Phlox family.

Gravendeel B, Smithson A, Slik FJ, Schuiteman A. 2004. Epiphytism and pollinator specialization: drivers for orchid diversity? *Philos Trans R Soc Lond B Biol Sci* **359**(1450): 1523-1535.

Harder LD, Johnson SD. 2009. Darwin's beautiful contrivances: evolutionary and functional evidence for floral adaptation. *New Phytol* **183**(3): 530-545.

Harvey PH, Pagel MD. 1991. *The comparative method in evolutionary biology*: Oxford university press Oxford.

Hermann K, Klahre U, Moser M, Sheehan H, Mandel T, Kuhlemeier C. 2013. Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in Petunia. *Curr Biol* **23**(10): 873-877.

Hermann K, Klahre U, Venail J, Brandenburg A, Kuhlemeier C. 2015. The genetics of reproductive organ morphology in two Petunia species with contrasting pollination syndromes. *Planta* **241**(5): 1241-1254.

Hillis DM, Moritz C, Mable BK, Olmstead RG. 1996. *Molecular systematics*: Sinauer Associates Sunderland, MA.

Hodges SA. 1997. Floral nectar spurs and diversification. *International Journal of Plant Sciences*: S81-S88.

Hodges SA, Arnold ML. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London B: Biological Sciences* **262**(1365): 343-348.

Hopkins R, Levin DA, Rausher MD. 2011. Molecular signatures of selection on reproductive character displacement of flower color in Phlox drummondii. *Evolution* **66**(2): 469-485.

Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic Mapping of Morphological Characters. *Systematic Biology* **52**(2): 131-158.

Kay KM, Reeves PA, Olmstead RG, Schemske DW. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *American Journal of Botany* **92**(11): 1899-1910.

Kay KM, Sargent RD. 2009. The Role of Animal Pollination in Plant Speciation: Integrating Ecology, Geography, and Genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**(1): 637-656.

Kay KM, Voelckel C, Yang JY, Hufford KM, Kaska DD, Hodges SA. 2006. Floral characters and species diversification. *Ecology and evolution of flowers*: 311-325.

Klahre U, Gurba A, Hermann K, Saxenhofer M, Bossolini E, Guerin PM, Kuhlemeier C. 2011. Pollinator choice in Petunia depends on two major genetic Loci for floral scent production. *Curr Biol* **21**(9): 730-739.

Lulin H, Xiao Y, Pei S, Wen T, Shangqin H. 2012. The first Illumina-based de novo transcriptome sequencing and analysis of safflower flowers. *PLoS One* **7**(6): e38653.

Maddison WP. 2006. Confounding Asymmetries in Evolutionary Diversification and Character Change. *Evolution* **60**(8): 1743.

Maddison WP, FitzJohn RG. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* **64**(1): 127-136.

Maddison WP, Midford PE, Otto SP. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* **56**(5): 701-710.

Martén-Rodríguez S, Almarales-Castro A, Fenster CB. 2009. Evaluation of pollination syndromes in Antillean Gesneriaceae: evidence for bat, hummingbird and generalized flowers. *Journal of Ecology* **97**(2): 348-359.

Marten-Rodriguez S, Fenster CB, Agnarsson I, Skog LE, Zimmer EA. 2010. Evolutionary breakdown of pollination specialization in a Caribbean plant radiation. *New Phytol* **188**(2): 403-417.

Martin A, Orgogozo V. 2013. The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**(5): 1235-1250.

Michez D, Vanderplanck M, Engel MS, Patiny S 2012. Fossil bees and their plant associates. In: Patiny S ed. *Evolution of plant-pollinator relationships*, 103-164.

Möller M, Clark J. 2013. The state of molecular studies in the family Gesneriaceae: a review. *Selbyana* **31**(2): 95-125.

Mora MM, Clark JL. 2016. Molecular Phylogeny of the Neotropical Genus *Paradrymonia* (Gesneriaceae), Reexamination of Generic Concepts and the Resurrection of *Trichodrymonia* and *Centrosolenia*. *Systematic Botany* **41**(1): 82-104.

Muchhala N, Johnsen S, Smith SD. 2014. Competition for hummingbird pollination shapes flower color variation in Andean solanaceae. *Evolution* **68**(8): 2275-2286.

Muchhala N, Potts MD. 2007. Character displacement among bat-pollinated flowers of the genus *Burmeistera*: analysis of mechanism, process and pattern. *Proc Biol Sci* **274**(1626): 2731-2737.

Nee S. 2001. Inferring speciation rates from phylogenies. *Evolution* **55**(4): 661-668.

Nee S. 2006. Birth-Death Models in Macroevolution. *Annual Review of Ecology, Evolution, and Systematics* **37**(1): 1-17.

O'Meara BC, Beaulieu JM. 2016. Past, future, and present of state-dependent models of diversification. *Am J Bot* **103**(5): 792-795.

Ollerton J, Alarcon R, Waser NM, Price MV, Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. *Ann Bot* **103**(9): 1471-1480.

Ollerton J, Winfree R, Tarrant S. 2011. How many flowering plants are pollinated by animals? *Oikos* **120**(3): 321-326.

Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society of London B: Biological Sciences* **255**(1342): 37-45.

Pellissier L, Moreira X, Danner H, Serrano-Serrano M, Salamin N, Dam NM, Rasmann S. 2016. The simultaneous inducibility of phytochemicals related to plant direct and indirect defences against herbivores is stronger at low elevation. *Journal of Ecology*.

Pennisi E. 2005. What determines species diversity? *Science* **309**(5731): 90.

Perret M, Chautems A, De Araujo AO, Salamin N. 2013. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. *Botanical Journal of the Linnean Society* **171**(1): 61-79.

Perret M, Chautems A, Spichiger R, Barraclough TG, Savolainen V. 2007. The geographical pattern of speciation and floral diversification in the neotropics: the tribe Sinningieae (Gesneriaceae) as a case study. *Evolution* **61**(7): 1641-1660.

Perret M, Chautems A, Spichiger R, Kite G, Savolainen V. 2003. Systematics and evolution of tribe Sinningieae (Gesneriaceae): evidence from phylogenetic analyses of six plastid DNA regions and nuclear ncpGS. *American Journal of Botany* **90**(3): 445-460.

Perret M, Chautems A, Spichiger R, Peixoto M, Savolainen V. 2001. Nectar sugar composition in relation to pollination syndromes in Sinningieae (Gesneriaceae). *Annals of Botany* **87**(2): 267-273.

Preston JC, Hileman LC, Cubas P. 2011. Reduce, reuse, and recycle: developmental evolution of trait diversification. *Am J Bot* **98**(3): 397-403.

Ramsey J, Bradshaw H, Schemske DW. 2003. Components of reproductive isolation between the monkeyflowers Mimulus lewisii and M. cardinalis (Phrymaceae). *Evolution* **57**(7): 1520-1534.

Ricklefs RE. 2007. Estimating diversification rates from phylogenetic information. *Trends Ecol Evol* **22**(11): 601-610.

Roalson EH, Skog LE, Zimmer EA. 2008. Untangling Gloxinieae (Gesneriaceae). II. Reconstructing Biogeographic Patterns and Estimating Divergence Times Among New World Continental and Island Lineages. *Systematic Botany* **33**(1): 159-175.

Sargent RD. 2004. Floral symmetry affects speciation rates in angiosperms. *Proc Biol Sci* **271**(1539): 603-608.

Savolainen O, Lascoux M, Merila J. 2013. Ecological genomics of local adaptation. *Nat Rev Genet* **14**(11): 807-820.

Schemske DW, Bradshaw H. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). *Proceedings of the National Academy of Sciences* **96**(21): 11910-11915.

Schiestl FP, Schluter PM. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annu Rev Entomol* **54**: 425-446.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS One* **8**(5): e64621.

Sheehan H, Moser M, Klahre U, Esfeld K, Dell'Olivo A, Mandel T, Metzger S, Vandenbussche M, Freitas L, Kuhlemeier C. 2016. MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. *Nat Genet* **48**(2): 159-166.

Singh VK, Garg R, Jain M. 2013. A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol J* **11**(6): 691-701.

Smith DS. 2010. Using phylogenetics to detect pollinator-mediated floral evolution. *New Phytol* **188**(2): 354-363.

Smith JF. 2000. A phylogenetic analysis of tribes Beslerieae and Napeantheae (Gesneriaceae) and evolution of fruit types: parsimony and maximum likelihood analyses of ndhF sequences. *Systematic Botany* **25**(1): 72-81.

Smith JF. 2001. High species diversity in fleshy-fruited tropical understory plants. *The American Naturalist* **157**(6): 646-653.

Smith JF, Clark JL. 2013. Molecular Phylogenetic Analyses Reveal Undiscovered Monospecific Genera in the tribe Episcieae (Gesneriaceae). *Systematic Botany* **38**(2): 451-463.

Smith SD, Rausher MD. 2011. Gene loss and parallel evolution contribute to species difference in flower color. *Mol Biol Evol* **28**(10): 2799-2810.

Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a flower color transition. *Mol Biol Evol* **30**(3): 602-612.

Stadler T. 2013. How can we improve accuracy of macroevolutionary rate estimates? *Syst Biol* **62**(2): 321-329.

Strauss SY, Whittall JB. 2006. Non-pollinator agents of selection on floral traits. *Ecology and evolution of flowers*: 120-138.

Stuurman J, Hoballah ME, Broger L, Moore J, Basten C, Kuhlemeier C. 2004. Dissection of floral pollination syndromes in Petunia. *Genetics* **168**(3): 1585-1599.

Thomson James D, Wilson P. 2008. Explaining Evolutionary Shifts between Bee and Hummingbird Pollination: Convergence, Divergence, and Directionality. *International Journal of Plant Sciences* **169**(1): 23-38.

Todd EV, Black MA, Gemmell NJ. 2016. The power and promise of RNA-seq in ecology and evolution. *Mol Ecol* **25**(6): 1224-1241.

Vamosi JC, Vamosi SM. 2011. Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. *Am J Bot* **98**(3): 460-471.

van der Niet T, Johnson SD. 2012. Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends Ecol Evol* **27**(6): 353-361.

Van der Niet T, Peakall R, Johnson SD. 2014. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Ann Bot* **113**(2): 199-211.

Weber A, Clark JL, Möller M. 2013. A new formal classification of Gesneriaceae. *Selbyana* **31**(2): 68-94.

Wessinger CA, Hileman LC, Rausher MD. 2014. Identification of major quantitative trait loci underlying floral pollination syndrome divergence in Penstemon. *Philos Trans R Soc Lond B Biol Sci* **369**(1648).

Wessinger CA, Rausher MD. 2014. Predictability and irreversibility of genetic changes associated with flower color evolution in Penstemon barbatus. *Evolution* **68**(4): 1058-1070.

Whittall JB, Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* **447**(7145): 706-709.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in Aquilegia. *Mol Ecol* **15**(14): 4645-4657.

Wilson P, Castellanos MC, Wolfe AD, Thomson JD 2006. Shifts between bee and bird pollination in Penstemons. *Plant–pollinator interactions: from specialization to generalization*, 47-68.

Chapter 1

Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae

From the article: **Serrano-Serrano ML,** Perret M, Guignard M, Chautems A, Silvestro D, Salamin N. **2015**. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evolutionary Biology* 15: 247. DOI: 10.1186/s12862-015-0527-6

Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae

Martha Liliana Serrano-Serrano[1,2], Mathieu Perret[3], Maïté Guignard[1], Alain Chautems[3], Daniele Silvestro[1,2,4] and Nicolas Salamin[1,2]

[1]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[2]Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland
[3]Conservatoire et Jardin botaniques de la Ville de Genève and Laboratory of Plant Systematics and Biodiversity, University of Geneva, Chemin de l'Impératrice, 1, 1292 Chambésy, Geneva, Switzerland
[4]Department of Plant and Environmental Sciences, University of Gothenburg, Carl Skottsbergs gata 22B, 413 19 Gothenburg, Sweden.

Summary

Background: Major factors influencing the phenotypic diversity of a lineage can be recognized by characterizing the extent and mode of trait evolution between related species. Here, we compared the evolutionary dynamics of traits associated with floral morphology and climatic preferences in a clade composed of the genera *Codonanthopsis*, *Codonanthe* and *Nematanthus* (Gesneriaceae). To test the mode and specific components that lead to phenotypic diversity in this group, we performed a Bayesian phylogenetic analysis of combined nuclear and plastid DNA sequences and model the evolution of quantitative traits related to flower shape and size and to climatic preferences. We propose an alternative approach to display graphically the complex dynamics of trait evolution along a phylogenetic tree using a wide range of evolutionary scenarios.

Results: Our results demonstrated heterogeneous trait evolution. Floral shapes displaced into separate regimes selected by the different pollinator types (hummingbirds versus insects), while floral size underwent a clade-specific evolution. Rates of evolution were higher for the clade that is hummingbird pollinated and experienced flower resupination, compared with species pollinated by bees, suggesting a relevant role of plant-pollinator interactions in lowland rainforest. The evolution of temperature preferences is best explained by a model with distinct selective regimes between the Brazilian Atlantic Forest and the other biomes, whereas differentiation along the precipitation axis was characterized by higher rates, compared with temperature, and no regime or clade-specific patterns.

Conclusions: Our study shows different selective regimes and clade-specific patterns in the evolution of morphological and climatic components during the diversification of Neotropical species. Our new graphical visualization tool allowed the representation of trait trajectories under parameter-rich models, thus contributing to a better understanding of complex evolutionary dynamics.

Keywords: Brazilian Atlantic forest, Hummingbird pollination, Traitgram, Resupination, Pollination syndrome, Trait evolution, Comparative methods

Background

Throughout the evolutionary process, lineages may experience divergent modifications of their phenotype and genome that culminate with the establishment of separate species. Modeling the evolution of species traits can help to elucidate the likely sequence of diversification events that lead to phenotypically diverse groups of species [1]. Traits related to different niche axes are expected to follow different evolutionary trajectories that may reflect different selection pressures, genetic constraints or stages of diversification determining the order in which the different ecological axes are partitioned during species divergence [2]. For example, patterns of trait divergence during the diversification of live oaks (*Ceanothus*) in California suggested that traits related to local scale coexistence show an early divergence in the group, while traits related to large scale habitat display a later or throughout differentiation [3]. Although theoretical work supports similar scenarios [1] empirical support for this model in other plant groups and across different types of traits still needs to be evaluated.

Hypotheses about the ordering of trait divergence during the evolution of a lineage can be complemented by investigating the mode and tempo of trait diversification within lineages [4]. For instance, initially rapid morphological evolution followed by relative stasis [5] could be the result of new ecological opportunities accompanied by density-dependent slowdowns in species diversification [6,7]. To explore this process, trait evolution can be reconstructed along the branches of phylogenetic trees to detect heterogeneity in evolutionary rates through time, across lineages or in relation to discrete characters [8-10]. Furthermore, Ornstein-Uhlenbeck (OU) models can be used to describe bounded phenotypic evolution, where single or multiple selective regimes pull phenotypes towards optimum values [11]. In plants, these models have helped to understand the evolutionary dynamics of flower morphology [12] and climatic niche [13]. Multiple studies have identified heterogeneous rates of evolution across climate dimensions in specific clades [14] and, at a larger scale, rates of niche evolution within major groups of angiosperms that are dependent on the type of growth form [15]. The possibility of testing multiple models to reveal complex patterns of trait evolution during species diversification is an important advantage to understand the dynamics of trait evolution and differential evolution among traits [7, 16, 17]. However, the fit between the current models and the real evolutionary processes is widely discussed [5, 18], and the power for selecting models depends on the number of taxa, the shape of the phylogeny, and the presence of measurements errors [19].

In this study, we investigate the evolutionary history of floral morphology and climatic preferences in a clade of epiphytic plants belonging to the genera *Codonanthopsis*, *Codonanthe*, and *Nematanthus* (hereafter referred to as the CCN clade) of the Gesneriaceae family. This group provides an excellent opportunity to compare patterns of evolutionary diversification between these niche axes. CCN clade exhibits a remarkable floral diversity in shape, size and orientation reflecting potential adaptation to different pollinators including bees and various hummingbirds [20-25]. Furthermore, CCN clade is widely distributed throughout most Neotropical rainforest but present in higher species richness and level of range overlap in the Brazilian Atlantic Forest (BAF) [26, 27]. Understanding how these morphological or climatic axes of niche differentiation have evolved in this plant group could shed light on the way speciation processes are building Neotropical biodiversity. First, we test if traits related to flower shape and size better fits a pollinator shift model involving transitions between adaptive peaks defined by pollinator morphology and behavior [12], or if flowers have diversified regardless of the pollinator type. Second, we determined if the evolution of climatic preferences is best explained by a model with distinct ecological optima [13] or a model with more labile evolution of climatic preferences among closely related species [14]. To address these questions we first infer phylogenetic relationships among the species using multi-gene DNA sequences. We quantify the floral morphology and climatic space occupied by the group and, examine the tempo and mode of evolution of different traits in the CCN group using current models of trait evolution. We finally develop a new approach to visualize the estimated trait evolution by proposing an alternative way to incorporate information from complex models.

Our results suggest that phenotypic evolution of this group is described by a variety of processes with different mode, time and lineage-specific effects. A new visualization of complex models of trait evolution further allow a better understanding of the particular processes at play in this group of Neotropical plants.

Methods

Taxonomic sampling

Taxonomic sampling included 46 out of the 52 species in the group, as well as 13 outgroup species. Six molecular makers, two nuclear (ITS and *ncpGS*) and four plastic regions (*atpB-rbcL* spacer, *rpl16* intron, *rps16* intron, *trnL-trnF* spacer) were sequenced and aligned for a final DNA matrix of 4.484 bp. We reconstructed phylogenetic relationships and relative divergence times using MrBayes and BEAST [28, 29]. Best fitting nucleotide substitution models were estimated with the *phymltest* function in ape R package [30]. Log-normal uncorrelated relaxed clock and the Yule speciation priors were set for the analyses. We used a maximum clade credibility (MCC) tree and a sample of high posterior probability trees from the BEAST results for later analyses. Finally, we examined the evolution of three binary traits (geographic distribution, pollination syndromes, and floral orientation) by reconstructing their ancestral states in the R package corHMM [31]. Detailed description of the molecular dataset, phylogenetic reconstructions, and ancestral state estimation are provided in the Appendix S1.

Morphometric data

Thirteen quantitative traits representing different aspects of the floral shape and size (Figure S1) were measured for 38 species out of the 46 included in the phylogenetic analysis (Table S3). Measurements were obtained from 2 to 15 flowers (average = 5) collected from wild individuals or cultivated plants at the Botanical Garden of Geneva (Switzerland). Collection permits were granted by the CNPq in Brazil (CMC 038/03) and the ANAM in Panama (SC/P-43-10). Floral material was not available for *Codonanthopsis dissimulata*, five species of *Codonanthe* and two *Nematanthus* (*C. crassifolia*, *C. calcarata*, *C. gibbosa*, *C. erubescens*, *C. luteola*, *N. kautskyi*, and *N. lanceolatus).* However, original descriptions of the species and photographic material available at Mauro Peixoto website (www.brazilplants.com) and at the Gesneriaceae Image Library (http://gesneriads.ua.edu/image-library/) indicates that the missing species do not represent exceptional morpho-types of the group. Thus, we can postulate that our quantitative measurements are representative of the morphological diversity of each clade and that we do not miss important variation because of the species lacking morphological data. Species positions in morphological space was quantified with a principal component analysis (PCA) using the R package *Ade4* [32] based on the covariance matrix of mean values for each species. We used non-transformed data, but log-transformed PCA patterns were also examined and led to a very similar morphospace (see figure S5).

Species climatic preferences

Climatic parameters for each species of the CCN clade were estimated from occurrence data and layers for climatic data. Locality descriptions were derived from the labels on specimens examined in more than 50 herbaria (Table S4). Georeferenced coordinates were generated for all localities that could be attributed to a precise geographic entity. We completed the dataset for species occurring outside Brazil with additional georeferenced specimens retrieved from GBIF (data.gbif.org, 2012-02-06). Occurrence data for *C. corniculata*, *C. elegans* and *N. serpens* were not included in the analysis due to limited number of herbarium material and uncertainty in their native distribution. A total of 2,240 occurrence points remained after manual checking and removal of duplicated points with a median of 15 occurrences per species. Climatic data (elevation and 19 bioclimatic variables) were extracted directly from Bioclim environmental layers [33] on a 30 arc-second

resolution grid (~1 km^2 at the equator). Occurrence data and their associated climatic information extracted from the Bioclim layers can be found at Dryad (http://dx.doi.org/10.5061/dryad.m871c). We used these climatic parameters to represent the species distribution along climatic gradients. We denoted this climatic space as the species preferences, although these preferences may be limited by interactions with other species, historical factors or dispersal limitation [34]. To explore the relative position of each species in the climatic space of the CCN clade, we performed a PCA using outlying mean index ordination (OMI) [35], which assigns a mean position of each species on the climatic space, as implemented in *Ade4* package in R [32]. Values for each of the 19 Bioclim variables, plus altitude, were used for the ordination.

Models of continuous trait evolution

We examined the patterns of trait evolution by using the MCC tree and multiple models, which span from a single Brownian motion rate of evolution (BM with single σ^2), BM with variable rate through time (decreasing or increasing σ^2) including early burst (EB) to Ornstein-Uhlenbeck models. All these models were fitted using the *fitContinuous* function in geiger R package [36]. The Ornstein-Uhlenbeck models were fitted using either a single selective regime (one single σ^2, selection strength α and optimum parameter θ) or multi-regime processes. For the latter case, we tested four different OU models using the *OUwie* function in the OUwie package [11] (Table S6 for acronyms), each with two different regime categories: pollination syndromes and geographic distribution. These categories were treated as binary characters, following the reconstructions described in Appendix S1. Delta-AIC (ΔAIC) and Akaike weights (ω) were calculated for model comparisons. Furthermore, Blomberg's K [37] measure of phylogenetic signal was estimated using phytools R package [38]. The outgroup species were pruned from the trees in all the morphology and climatic preferences analyses.

In addition to the models of trait evolution, a multi-rate BM model was tested using the *rjmcmc.bm* functions in the Geiger package in R [36]. This flexible method aims to identify changes on rates of continuous trait evolution among lineages. The analyses were performed on 100 phylogenetic trees randomly sampled from the BEAST posterior distribution and the MCMC was run for 100,000 generations, sampling every 100 generations and excluding the first 25% for burn-in. Each run provides posterior distributions of branch-wise rate estimates and probabilities of rate shifts. We compared the fit of the multi-rates model against the fit of alternative BM and OU models by comparing their AIC value. The multi-rates BM analysis used reversible jump MCMC (rjMCMC) and we estimated the AIC values for each model in two ways. First, we directly took the best likelihood value sampled over the rjMCMC samples and calculate the corresponding AIC value based on the number of parameters of the model. Second, we mapped the set of branches at each rate category with the *make.era.map* function, and used the non-censored approach implemented in the *brownie.lite* function (phytools R package) [38] to produce the maximum likelihood estimate and calculate the AIC values. A bias towards parameter-rich models can occur during the model selection process if measurement error is not considered [19]. We performed additional model comparisons to test for such effects by estimating the amount of measurement error present in our dataset. Such estimation is not possible for multi-rates BM model but we used the values estimated under single BM model estimation for these cases.

Visualization of continuous trait evolution

The parameters estimated under the best fitting model of trait evolution describe the process numerically, but a graphical visualization of trait evolutionary trajectories remains difficult to picture. Recent graphical methods (often referred to as "traitgrams") help to visualize phenotypic evolution by plotting a phylogenetic tree against trait values [39]. However, such methods are limited to the BM model, partly due to the difficulty of inferring ancestral states under more complex evolutionary models. Here, we implemented an alternative approach to display the dynamics of trait evolution along a phylogenetic tree. We achieved this by

forward-time simulations under complex models of trait evolution and analytical interpolation between ancestral states. The purpose of these reconstructions is to provide a graphical representation of the expected continuous evolution of phenotypes given a phylogeny and a set of parameters describing the evolutionary process. For each trait, we recorded the topological placement and magnitude of parameter changes across the tree, e.g. rates, selection strength and optima, depending on the model (Table 1). We then simulated 100 realizations of trait evolution under the optimized model and parameter values along the BEAST MCC tree (see Appendix S2 for more details). All forward simulations started by sampling a random number in a normal distribution with parameters estimated from a posterior distribution of root states obtained with the *fitContinuousMCMC* function in geiger [36]. Other parameters (σ^2, α, θ) were taken from the best fitting model.

We used the simulated trait values at the internal nodes and at the tips to plot traitgrams displaying the reconstructed trait evolution. In their standard implementation traitgrams connects the trait values between nodes by a straight line to draw the edges of a tree [39, 40], consistently with the expected anagenetic trait evolution under a constant rate BM model. However, under more complex models such as OU processes, the expected trait value $x(t)$ at a time t between two nodes of age $t_i < t < t_j$ and trait values $x(t_i)$ and $x(t_j)$ respectively, deviates from a straight line. This deviation occurs because the evolutionary trajectory of the trait under an OU model depends not only on the evolutionary rate, but also on the strength of selection and the relative distance from the optimum [41]. The expected anagenetic evolution of a trait under an OU model is described by the Ornstein-Uhlenbeck bridge (personal communication) [42] and can be obtained for any time t as:

$$x(t) = \theta + \frac{(x(t_i)-\theta)\sinh(\alpha(t_j-t))}{\sinh(\alpha(t_j-t_i))} + \frac{(x(t_i)-\theta)\sinh(\alpha(t-t_i))}{\sinh(\alpha(t_j-t_i))}, \quad \text{for} \quad t_i < t < t_j$$

Where θ is the optimum and α the strength of selection. The parameters θ and α can vary across the edges of a tree if the model includes different selective regimes.

We generated multiple realizations of the trait evolution under the different BM and OU models and parameter settings estimated for the four traits. These were combined to plot the 95% CI of the trait ranges through time (i.e., the minimum and maximum trait values across clades at any time t). These plots thus provide a graphical representation of how the range of potential trait values are expected to change through time, given a fixed tree topology and a complex model of trait evolution. Single realizations of the simulated process, resembling the conventional traitgrams, facilitate the understanding of regime- or clade-specific patterns. The method is implemented in an R script available at www.unil.ch/phylo/bioinformatics.

Results

Phylogenetic analyses

The tree topologies reconstructed from MrBayes and BEAST were congruent and generally highly supported (see figures S2 and S6). The BEAST analysis (Fig. 1, see Fig. S2 for tree with outgroups) indicated that the initial divergence in the CCN lineage occurred between the *Codonanthopsis* clade and the remaining taxa that are all endemic to the BAF. The *Codonanthopsis* clade is widely distributed in Central America, the Caribbean, the Andes and the Amazonian basin, but does not occur in the BAF except for the widespread *C. uleana*, which extends its range into the northern part of this biome. The BAF clade is composed of the two sister clades corresponding to the *Codonanthe s.s.* (seven species) and *Nematanthus* (27 species). Within *Nematanthus*, the sister species *N. australis* and *N. wettsteinii* first diverged from the remaining *Nematanthus*, which includes two sister clades identified as *Nematanthus A* (17 species) and *Nematanthus B* (7 species).

High posterior probabilities supported these clades (1.0), while ambiguous placements were observed for *C. mattos-silvae*, *N. kaustkyi* and *N. hirtellus* species. Ancestral reconstructions for geography, pollination syndrome and floral orientation using the preferred model (ER) showed few state changes, mainly involving complete clade transitions (Figure S3).

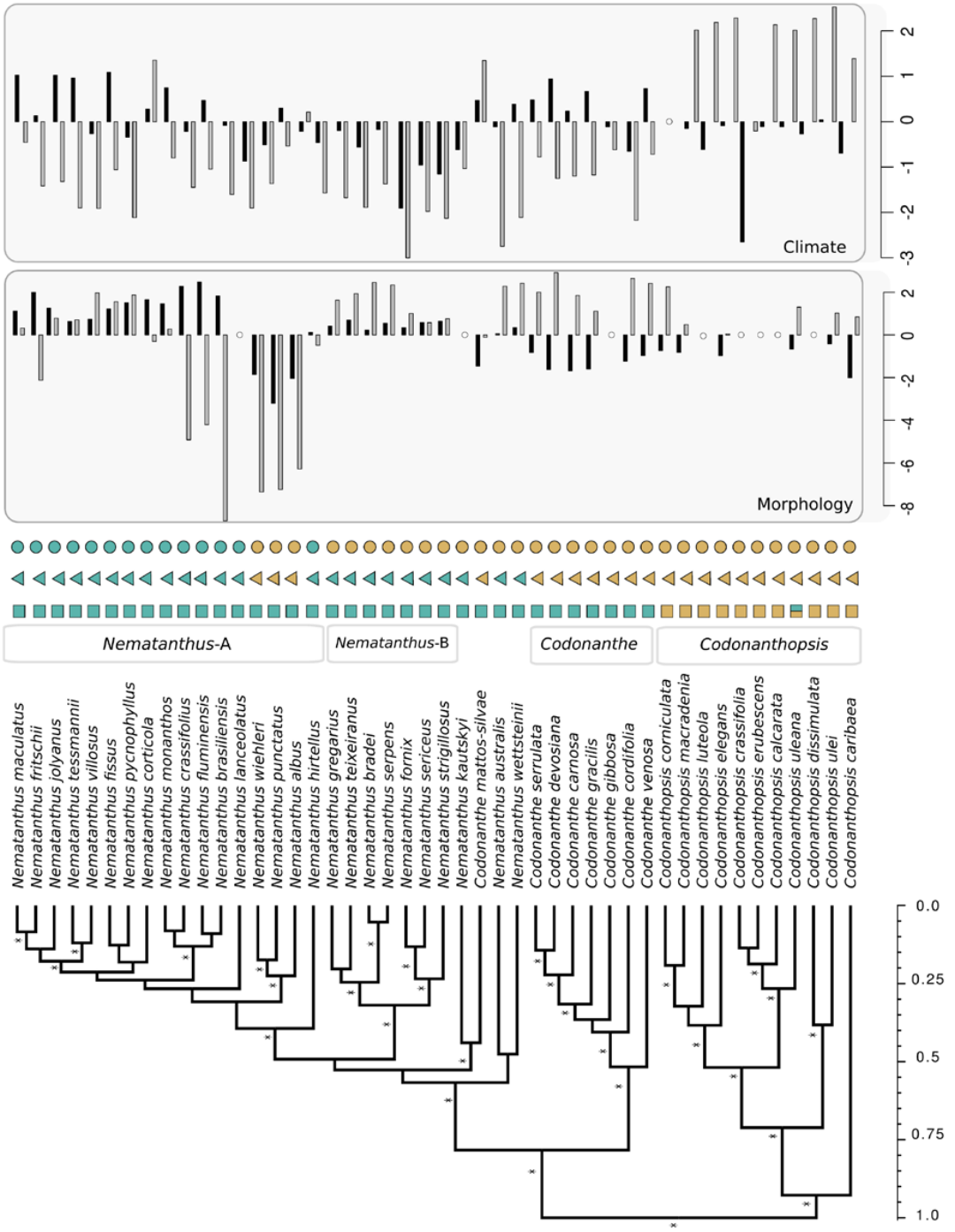


Figure 1. Maximum clade credibility tree. Asterisk on branches are Bayesian posterior probabilities >0.99. Subclades are indicated by gray-shaded boxes. Binary traits are indicated for each species on top of the tree. Geographical distribution: green square = BAF, yellow square = other biomes. Pollination syndromes: green triangle = hummingbird, yellow triangle = bee pollinated, and floral orientation: green circle = resupinate, yellow circle = non-resupinate. Principal component values for morphology and climate are enclosed in gray frame, upper panels. Gray and black vertical bars represent PC1 and PC2, respectively.

Floral morphology and climatic preferences space

The first and second axes of the PCA on the floral traits explained 85.1% of the variance (PC1 70.7%; PC2 14.4%, Figure S4a and Table S5). PC1 mainly reflected variation in flower size with loadings of same sign and approximately equal value for all measurements. Variation in flower size (PC1) was particularly extensive in the *Nematanthus*-A clade (Fig. S4a). PC2 had a positive loading for stamen, pistil and tube lengths, vertical diameter of corolla tube, and a negative loading for the diameters of limb and corolla opening and restriction before nectary chamber. PC2 therefore mainly represented variation in flower shape, with a positive value indicating tubular and narrowly opened corolla, while a negative value indicates more campanulate corolla with inserted stamen and broad limb and opening. All species with positive value of PC2 belong to the genus *Nematanthus.* Their flowers are pigmented in red, orange or yellow and match well the definition of the syndrome of hummingbird pollination (Fenster et al. 2004), as confirmed by field studies for 10 species (see Table S2). Species with negative PC2 values belong to the *Codonanthe, Codonanthopsis* and to a specific subclade within *Nematanthus*-A. Their flowers have several features traditionally associated with bee pollination such as the creamy corolla, brownish dots inside the tube (nectary guide), inserted stamen and developed inferior lobes forming a landing platform for insects. However, to our knowledge, no field studies have confirmed bee pollination for these species.

The first two PCs for climatic variables accounted for 69.1% of the variance (PC1 54.30%, PC2 14.84**%,** Fig. S4b)**.** PC1 reflected mainly the variation in temperature (Bio9, Bio11, Bio4 and Bio7) with positive values indicating warmer mean temperatures with low seasonal variation and negative values showing strong variability in temperature through the year. PC2 was mainly correlated with precipitation (Bio15 and Bio14). Positive PC2 values indicated high seasonal variability in precipitation regimes, while negative values showed high precipitation on the wettest month and low seasonality (see Fig. S4b for species climatic values). Climatic space indicated a clear separation between the *Codonanthopsis* species and the BAF lineages (*Codonanthe* plus *Nematanthus* clades), in agreement with their distinct geographical distribution.

Models of continuous trait evolution

The models of trait evolution indicated that distinct evolutionary processes have influenced trait divergence in the CCN group. Individual axes of floral morphology (size and shape) and climatic preferences (temperature and precipitation) variation seemed to have evolved independently. The incorporation of measurement error in our model comparisons did not lead to any bias towards parameter-rich models and the four phenotypic axes present no changes in the preferred models when accounting for it (Table S7).

The evolution of floral size (PC1) was best described by the multi-rates BM model with branch-specific rates (Table 1). The ΔAIC is very large suggesting that the alternative models poorly represent the evolution of the floral size. The posterior probabilities for the rates of evolution across the phylogeny supported one rate shift in this trait (Figure 3), which is associated with the origin of the *Nematanthus*-A clade (shift probability > 0.625). The increase in the estimated rates of evolution (σ^2) ranged from 3.049 to 136.076. This shift produced a deviation from a constant BM process, which is consistent with the observed low values of Blomberg's K statistic for PC1 (mean of 0.496 with 25% and 75% quantiles of 0.463 and 0.546, respectively). According to the parameter estimated by the best fitting model, the visualization of floral size evolution shows narrow trait ranges in the early stages of the diversification of the CCN group, followed by a large increase in trait ranges due to the single clade shift (Fig. 2a). From the simulated trait trajectories, and because the best model was based on BM, the increase in trait range can be symmetric (positive or negative axis). However, the empirical data suggested that only deviation towards bigger floral sizes occurred (negative loadings).

The evolution of the floral shape (PC2) was best explained by an OUMVA model with two regimes, which are defined by the pollination types. Each regime had different rates of evolution (σ^2), optima (θ) and strength of selection (α). Bee-pollinated species (negative loadings in the PC2) evolved in a constrained way, with a narrower dispersion from the optimum value, whereas hummingbird-pollinated species (positive loadings in PC2) explored a wider trait space during their evolution (Fig. 2b). Phylogenetic reconstruction indicated a single origin of hummingbird syndrome at the root of the *Nematanthus* clade and more recent reversal to bee syndrome in a clade of three species (*N. albus*, *N. punctatus*, and *N. wiehleri*, Fig. S3). Floral shape constraint agrees with the high Blomberg's K values obtained for the morphological PC2 (mean value of 1.323 with 25% and 75% quantiles of 1.122 and 1.547, respectively).

Table 1. Results of model fitting for the morphological and climate PC axes.

Models	AICc	ΔAIC	ω
Floral size – Morphology PC1			
Brownian Motion	190.9718	38.7133	0.0000
Ornstein-Uhlenbeck	187.7321	35.4736	0.0000
OU alternative[a]	172.9199	20.6614	0.0000
Early Burst (DC)	193.3461	41.0876	0.0000
Early burst (AC)	187.7321	35.4736	0.0000
Multiple rates	152.2585	0.0000	1.0000
Floral shape – Morphology PC2			
Brownian Motion	104.3669	24.8268	0.0000
Ornstein-Uhlenbeck	106.6641	27.1240	0.0000
OU alternative[a]	79.5401	0.0000	0.9996
Early Burst (DC)	106.7412	27.2011	0.0000
Early burst (AC)	106.6641	27.1240	0.0000
Multiple rates	95.2293	15.6892	0.0004
Mean and seasonality in temperature - Climate PC1			
Brownian Motion	134.5776	10.1545	0.0059
Ornstein-Uhlenbeck	135.6864	11.2633	0.0034
OU alternative[a]	124.4231	0.0000	0.9465
Early Burst (DC)	136.9015	12.4784	0.0018
Early burst (AC)	136.5797	12.1566	0.0022
Multiple rates	130.7428	6.3197	0.0402
Precipitation seasonality - Climate PC2			
Brownian Motion	115.0848	19.9517	0.0000
Ornstein-Uhlenbeck	100.5829	5.4498	0.0548
OU alternative[a]	100.5768	5.4437	0.0549
Early Burst (DC)	117.4086	22.2755	0.0000
Early burst (AC)	100.5829	5.4498	0.0548
Multiple rates	95.1331	0.0000	0.8355

[a]OU alternative corresponds to Ornstein–Uhlenbeck models with different numbers of parameters. Here, only the model with the best AIC value is reported, see the full set of OU models in summarized in the Additional file 1: Table S6. (AICc = corrected Akaike Information Criterion values, ΔAIC Delta AICc and ω = Akaike weights)

The evolution of climatic preferences also showed different dynamics among its components. The best model for temperature (climate PC1) was OUM, with different optima between the two geographic distributions (i.e., Central, northern South America and Amazonian basin versus Brazilian Atlantic forest), but equal rates of evolution and selection coefficient per regime. The trait space for temperature displayed a strongly bounded evolution with a slow rate of change (Fig. 2c). The extent of change in positive and negative loadings appeared to be symmetrical. Estimates of Blomberg's K statistic for climatic PC1 indicated a mean value of 0.477 with 25% and 75% quantiles of 0.376 and 0.541, respectively. Multi-rates BM model was the best fitted for precipitation seasonality (climatic PC2), but posterior evidence for a rate shift was weak. Only a minor increase in rates of climatic differentiation between the sister species *C. erubescens* and *C. crassifolia* was detected (see figure 4). The evolution of trait space for the climatic PC2 appeared as a constant increase of phenotypic space over time (Fig. 2d). This was consistent with the results in figure 4, showing that large shifts in the trait values are rare. Estimates of Blomberg's K values for climatic PC2 ranged from 0.357 and 0.459 (25% and 75% quantiles respectively) with a mean value of 0.408.

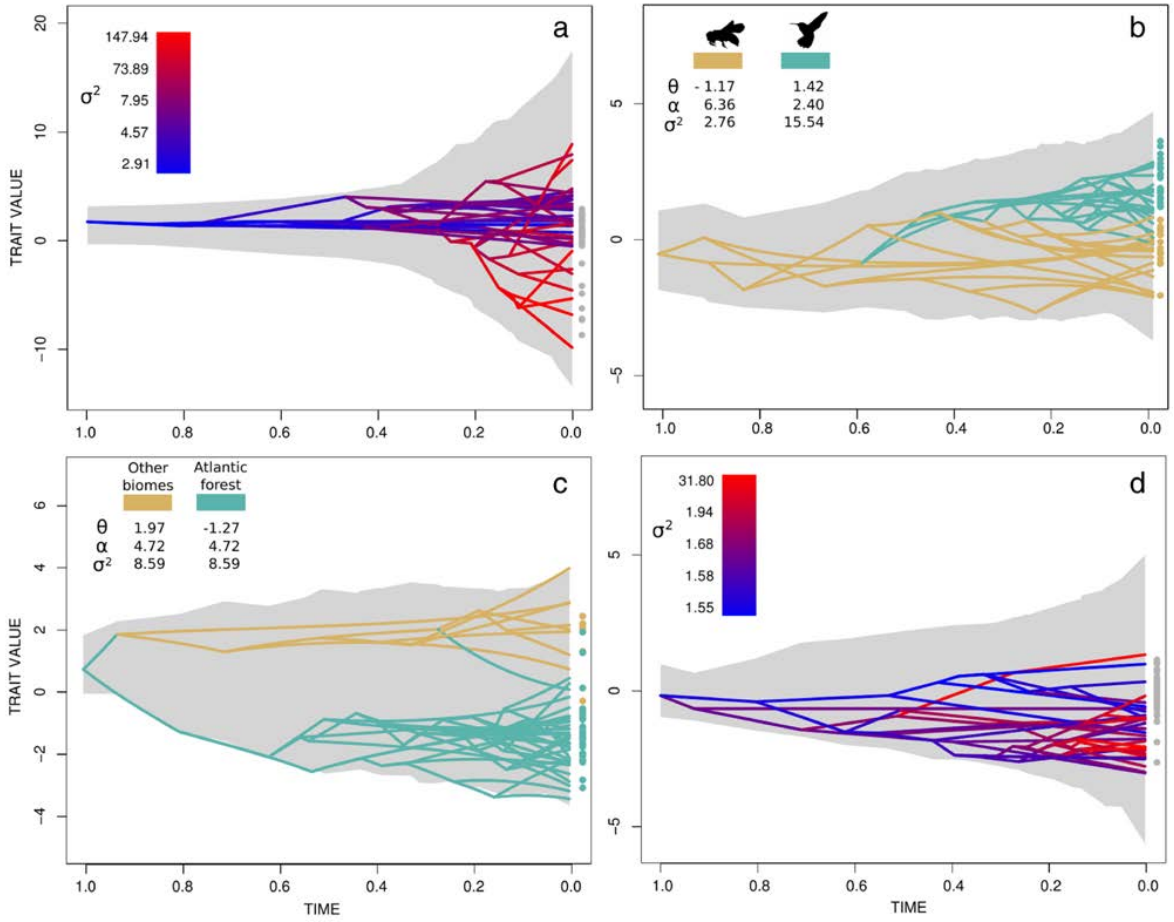


Figure 2. Simulated trait space and traitgrams (under specific models in table 1) for morphological and climatic traits. The Y-axis corresponds to the trait values for the species, and should not be confound with the variances between them, thus the gray-shaded area is the 95% CI of simulated trait ranges. Panel *a*, floral size (morphological PC1) with multiple Brownian motion model. Panel *b*, floral shape (morphological PC2) with regimes of the OUMVA model defined as bee and hummingbird pollinated species. Panel *c*, mean and seasonality in temperature (climatic PC1) with regimes in a OUM model defined as Atlantic forest and other biomes. Panel *d*, precipitation seasonality (climatic PC2) with multiple Brownian motion model. Colored scale in *a* and *d* correspond to branch-specific rates of trait evolution. Colors in *b* and *c* correspond to multiple regimes. Parameters θ, α and σ^2 correspond to the optimum, strength of selection and rates of evolution, respectively according to the model specification. Points at the right of each panel indicates the observed trait values in all species analyzed, yellow and green colors if regimes in the model.

Discussion

Testing the order and extent of trait divergence during the evolution of a clade helps to understand the relative importance of separate morphological and climatic trajectories, as well as the possible drivers of species diversification. We combine a phylogenetic analysis and multiple models of trait evolution with simulations under the selected models, in order to comprehensively understand these different trajectories of trait evolution in the CCN group. Our results suggest that phenotypic evolution of this group is described by a variety of processes with different mode, time and lineage-specific effects. A new visualization of complex models of trait evolution further allow a better understanding of the particular processes at play in this group of Neotropical plants.

Floral evolution dynamics

The inference of evolutionary models and estimation of plausible trait ranges for floral morphology revealed contrasting patterns during the evolutionary history of the CCN group. Floral size, represented by morphological PC1, has evolved in a complex fashion. The estimated trait range through time showed an initial period of narrow divergence, followed by a marked increase in trait ranges associated with the accelerated evolution of flower size within the clade *Nematanthus*-A (Fig. 2a). We did not detect evidence of a slowdown in the rate of evolution of the PC1, showing that divergence in floral size continues throughout the diversification of the *Nematanthus-A* lineage during the Miocene (24 Mya, 95% HPD 33.45 – 9.30 Mya [27]). This result contrasts with the classical model of adaptive radiation, where morphological evolution is initially rapid and then slows through time [5]. Our analyses were not aimed at investigating whether the CCN clade is a case of adaptive radiation, however the lack of a slowdown in the evolutionary trajectory of floral size suggest that morphological space is not yet filled.

The increase in rates of evolution of floral size detected at the base of the *Nematanthus*-A clade coincides with the evolution of floral resupination (see the placement of the rate shift in figure 3, and the most probable transition to resupinate flowers occurring in almost all species of the clade except for *N. albus, N. wiehleri* and *N. punctatus* in figure S3). A direct consequence of the evolution of flower resupination is a change of pollen placement on the body of the pollinator. In resupinate species belonging to the *Nematanthus*-A clade, pollen is primarily transported on different parts of the ventral side of hummingbirds [20, 23]. In contrast, non-resupinate *Nematanthus* place most of their pollen on the top of the bill [22]. Therefore, flower resupination and the associated shift of pollen deposition could have stimulated the diversification of floral size in *Nematanthus*-A clade by creating new opportunities for species coexistence while sharing pollinators as was shown in another community of hummingbird-pollinated plants [43]. Extending this analysis to other Gesneriaceae clades such as *Glossoloma* and *Crantzia* that independently evolved resupinate flowers pollinated by hummingbirds, would provide a mean to further test the positive effect of resupination on flower diversification [44, 45]. An additional feature of the *Nematanthus*-A clade is its contrasting altitudinal distribution compared with *Nematanthus*-B (median values of 524 m and 957 m, respectively). In the BAF, these two altitudinal levels have contrasting hummingbird assemblages, with lowland communities presenting a greater richness of hummingbirds and a corresponding higher heterogeneity in bill size (short-billed trochilines and long-billed hermits) compared with the highland sites [21, 22]. Indeed, *Nematanthus*-A species in lowland rainforest are visited by both Trochilinae (non-hermit) and long-billed hermit hummingbirds (*Phaethornis* and *Ramphodon*), whereas species from *Nematanthus-B* mainly rely on Trochilinae hummingbirds for pollination (see Table S2). We suggest that the interaction of *Nematanthus-A* species with broader range of hummingbird types and bill lengths in the lowland forest could also have provided more opportunities for flower size diversification in this clade.

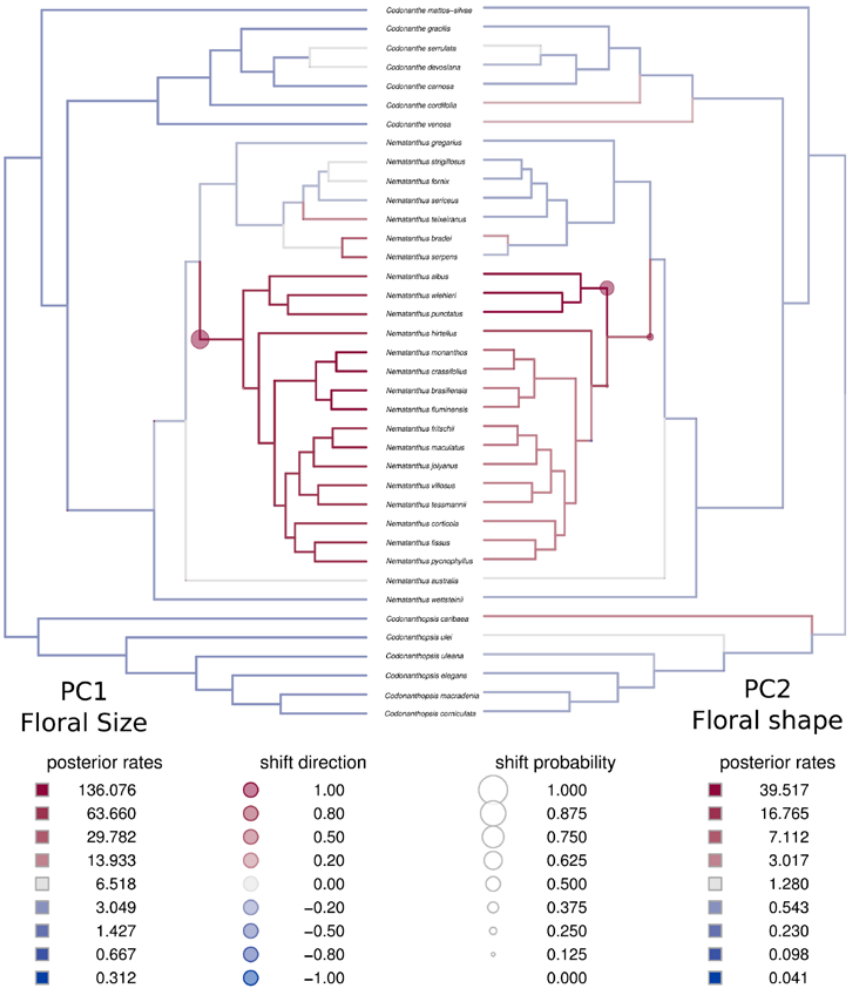


Figure 3. Posterior comparisons of rate of morphological trait evolution in the CCN group. Hue and size of circles at branches denote posterior support for a rate shift at the indicated branch. Larger and redder circles suggest higher posterior support for an upturn in evolutionary rate (see Eastman et al. 2011). Branches in the phylogeny are colored such that rates not deviant from the median are shaded gray; rates below (or above) the median are shaded blue (or red). Rates corresponding to each hue are indicated in the legend, as well as shift probabilities and directions.

Contrary to flower size, the evolution of the morphological PC2, floral shape, was preferentially supported by OU models that contained differential strength of selection, optima and rate parameters between hummingbird and bee pollination syndromes. The evolution of floral shape, from funnel-shaped corolla with expanded lobes in *Codonanthopsis* and *Codonanthe* clades to narrow-mouth and pouched corollas in most *Nematanthus* species could have been differentially constrained and maintained by functional groups of pollinators in agreement with the pollination syndrome concept [46]. Hummingbird-pollinated species showed a lower strength of selection and higher rates of trait change than species with a bee pollination syndrome (parameters in Fig. 2b), suggesting that hummingbirds might interact with a wider range of flower shapes than bees. The visualization of the floral shape trait space in this group provides a convenient tool that helps to better understand the dynamics of the OU process and the evolution of the independent optima.

Overall, our results suggest that the evolutionary trajectory of floral morphology in the CCN group may be constrained in shape (PC2), with possible evolutionary transitions from one functional group of pollinators to another following the pollinator shift model [12]. The transitions between these two phenotypic clusters may reflect selection to improve the interaction with better pollinators [46], and/or to avoid less efficient floral-pollinator associations [47, 48]. In comparison, variation in floral size (PC1) could be more related to character displacement and the establishment of mechanical isolation between co-occurring species sharing a same functional group of pollinators [49, 50]. This model of flower evolution based on competition for hummingbird pollination has been shown to generate phenotypic over-dispersion within communities of Andean Solanaceae [51]. Testing this prediction in the CCN clade would require to further investigate whether co-occurring species are more different in flower size than expected due to chance across different sites in the BAF.

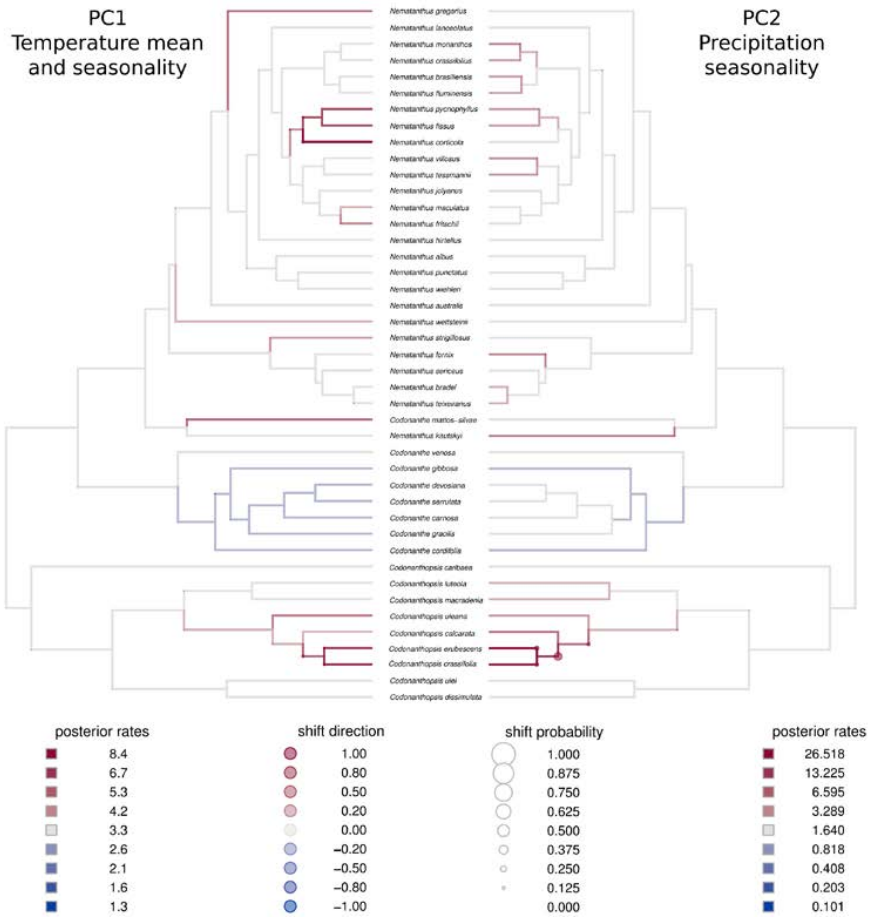


Figure 4. Posterior comparisons of rate of climatic preferences evolution in the CCN group. See caption of Figure 3.

Climatic evolution dynamics

The evolution of climatic preferences, represented in PC1 by temperature, is best explained by an OU model (Table 1). The different optima in the OUM model indicate the differentiation between BAF and other rainforests in the Neotropical region. However, models accounting for different rates of evolution and/or strength of selection are not preferred. This result potentially indicates that species preferences in temperature evolve at a similar pace, and that the strength of selection is comparable between groups inhabiting different regions. This pattern could reflect specific regional variation and dispersal limitation among the CCN lineages.

The graphical representation of the evolution of the temperature preference showed an early differentiation between the selective regimes with rare subsequent transitions from one biome to the other (Fig. 2c). In contrast, the multi-rate BM model best explained the precipitation preferences, which are represented by PC2. The estimated rates for this component were higher than the overall rates in PC1 (see parameters in fig. 2c and branch-specific estimations in fig. 4), suggesting that species preferences for precipitation seasonality might change more rapidly than temperature preferences which are mainly biome specific.

Several evolutionary studies have reported pronounced ecological niche differentiation, concentrated in particular lineages [13, 14, 52] or associated with distinct species traits [53, 54], which suggest an important role of climatic changes during speciation. Our results for climatic differentiation suggest an early biome separation of the CCN clades (BAF and the rest of the Neotropics), followed by a divergence along different local conditions of precipitation seasonality. Although the role of this climatic component in speciation would need to be further investigated, this result is consistent with previous studies showing that floristic turnover in the Atlantic forests is largely correlated with distance from the ocean and rainfall distribution patterns [55, 56] and that allopatric speciation could have been particularly frequent along this climatic gradient [57, 58].

Conclusions

Our investigation of the mode and tempo of trait evolution in the CCN clade provided evidence for a contrasting relevance of morphological and ecological divergences during species diversification. Two traits – flower shape and temperature preferences – were segregated into adaptive zones associated with different functional group of pollinators or biogeographic regions. First, floral shape evolution was constrained reflecting the selection to different functional groups of pollinators (i.e. hummingbirds vs insects). Second, divergence in temperature was linked with the colonization of the BAF biome at an early stage of the evolution of the CCN group. On the contrary, two other trait components – flower size and precipitation preferences – evolved at a higher rate, with no recent slowdown. Changes in floral size occurred mainly in an specific subclade including species with resupinate flowers and lowland distribution, whereas evolutionary changes of precipitation seasonality likely took place tree-wide and throughout the entire CCN diversification. The contrasting patterns between the constrained evolution of floral shape (i.e. pollination syndrome) and the throughout diversification of precipitation seasonality across time agree with the habitat first rule model proposed by Ackerly et al. (2006) suggesting early divergence of traits that allow species to co-occur (alpha niche) and a throughout diversification of traits defining macrohabitats (beta niche). We found however that both flower morphology and climatic preferences diversified along different axes that are better fitted by distinct models of evolution. Our new implementation for visualization allowed us to graphically represent trait evolutionary histories, using models that go beyond the simple BM and can potentially better capture complex evolutionary dynamics. This approach represents an enhancement of current methods to plot the phenotypic space through time, with the simulation part being of potential use as a predictive tool for measuring the power and fit of alternative models. Finally, our study calls for a broader phylogenetic scale analysis to unravel the mechanisms driving such evolutionary processes and their potential effect on the remarkable species richness in the Neotropics.

Additional material

1. R script available at http://www2.unil.ch/ phylo/files/software/plot_traitgram_serranoetal15.R
2. Occurrence data and their associated climatic information extracted from the Bioclim layers can be found at http://www2.unil.ch/phylo/ files/serranoetal15_bioclim.xls
3. Additional files in https://static-content.springer.com/esm/art%3A10.1186%2Fs12862-015-0527-6/MediaObjects/12862_2015_527_MOESM1_ESM.pdf

Authors' contributions

MLSS and NS conceived the study. MP, AC and MG participated significantly in the production of the data. MLSS performed lab work and analyses with the help of DS. NS and MP supervised MLSS, coordinated the project and helped to draft the manuscript. All authors have read and approved the manuscript.

Acknowledgements

We thank Régine Niba and Fadil Avdija for their contribution to the laboratory work; Pierre Mattille, Yvonne Menneret, Marina Magnette, and Matthieu Grillet for the propagation and maintenance of the Gesneriaceae collection at Geneva. Collection permits were granted by the CNPq in Brazil (CMC 038/03) and the ANAM in Panama (SC/P-43-10). We thank Mauro Peixoto for the kind supply of the photographic and plant material. We also thank Linda Dib, Anna Kostikova, Sacha Laurent, Glenn Litsios, Hervé Sauquet and Sylvain Corlay for the helpful discussions. We thank Simon Joly and an anonymous reviewer for insightful comments that greatly improved the manuscript. All analyses were run at the High-performance Computing Center (Vital-IT) from the Swiss Institute of Bioinformatics. This study was funded by the Faculty of Biology and Medicine at the University of Lausanne and the Swiss National Science foundation (grant number CRSII3_147630). DS is funded by the Wenner-Gren Foundation and Carl Tryggers stiftelse (Sweden).

References

1. Gravilets S. and Losos, J. Adaptive radiation: Contrasting theory with data. Science. 2009. 323: 732-737.

2. Schluter, D. The ecology of adaptive radiation. Oxford University Press, Oxford, UK. 2000

3. Ackerly, D., D. Schwilk, and C. O. Webb. Niche evolution and adaptive radiation: testing the order of trait divergence. Ecology. 2006; 87:50–61.

4. Glor, R. E. Phylogenetic Insights on Adaptive Radiation. Annual Review of Ecology, Evolution, and Systematics. 2010; 41:251–270.

5. Harmon, L. J., J. B. Losos, T. J. Davies, R. G. Gillespie, J. L. Gittleman, W. B. Jennings, K. H. Kozak, M. A. Mcpeek, F. Moreno-roark, T. J. Near, A. Purvis, R. E. Ricklefs, D. Schluter, J. A. S. Ii, O. Seehausen, B. L. Sidlauskas, O. Torres-carvajal, J. T. Weir, and A. Ø. Mooers. Early bursts of body size and shape evolution are rare in comparative data. Evolution. 2010; 2385–2396.

6. Mahler, D. L., L. J. Revell, R. E. Glor, and J. B. Losos. Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. Evolution. 2010; 64:2731–45.

7. Weir, J. T., and S. Mursleen. Diversity-dependent cladogenesis and trait evolution in the adaptive radiation of the auks (Aves: Alcidae). Evolution. 2012; 67:403–416.

8. O'Meara, B. C., C. Ané, M. J. Sanderson, and P. C. Wainwright. Testing for different rates of continuous trait evolution using likelihood. Evolution. 2006; 60:922–933.

9. Eastman, J. M., M. E. Alfaro, P. Joyce, A. L. Hipp, and L. J. Harmon. A novel comparative method for identifying shifts in the rate of character evolution on trees. Evolution. 2011; 3578–3589.

10. Revell, L. J., D. L. Mahler, P. R. Peres-neto, and B. D. Redelings. A new phylogenetic method for identifying exceptional phenotypic diversification. Evolution. 2011; 66 (1). 135–146.

11. Beaulieu, J. M., D.-C. Jhwueng, C. Boettiger, and B. C. O'Meara. Modeling Stabilizing Selection: Expanding the Ornstein-Uhlenbeck Model of Adaptive Evolution. Evolution. 2012; 66:2369–2383.

12. Whittall J.B. and Hodges S.A. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. Nature Letters. 2007. 447: 706-710.

13. Schnitzler, J., C. H. Graham, C. F. Dormann, K. Schiffers, and H. Peter Linder. Climatic niche evolution and species diversification in the Cape flora, South Africa. Journal of Biogeography. 2012; 39:2201–2211.

14. Evans, M. E. K., S. A. Smith, R. S. Flynn, and M. J. Donoghue. Climate , niche evolution , and diversification of the "Bird-Cage" evening primroses (*Oenothera*, Sections Anogra and Kleinia). The American Naturalist. 2009; 173:225–240.

15. Smith, S. a, and J. M. Beaulieu. Life history influences rates of climatic niche evolution in flowering plants. Proceedings of the Royal Society of London. Series B. 2009; 276:4345–52.

16. Kozak, K. H., and J. J. Wiens. Accelerated rates of climatic-niche evolution underlie rapid species diversification. Ecology Letters. 2010; 13:1378–1389.

17. Hertz, P. E., Y. Arima, A. Harrison, R. B. Huey, J. B. Losos, and R. E. Glor. Asynchronous evolution of physiology and morphology in Anolis lizards. Evolution. 2013; 67:2101–13.

18. Ingram, T., L. J. Harmon, and J. B. Shurin. When should we expect early bursts of trait evolution in comparative data? Predictions from an evolutionary food web model. Journal of Evolutionary Biology. 2012; 25:1902–1910.

19. Silvestro, D., A. Kostikova, G. Litsios, P.B. Pearman, and N. Salamin. Measurement errors should always be incorporated in phylogenetic comparative analysis. Methods in Ecology and Evolution. 2015; 2041-210X.

20. Franco A.L.M. and Buzato S. Biologia floral de *Nematanthus fritschii* (Gesneriaceae). Rev. Brasil. Biol. 1992; 52:661-666.

21. Sazima, I., S. Buzato, and M. Sazima. An assemblage of hummingbird-pollinated flowers in a montane forest in southeastern Brazil. Botanica Acta. 1996; 109:149–160.

22. Buzato, S., M. Sazima, and I. Sazima. Hummingbird-pollinated floras at three Atlantic forest sites. Biotropica. 2000; 32:824– 841.

23. Sanmartin-Gajardo, I. & Vianna J.R.S. Pollination of *Nematanthus brasiliensis:* An epiphytic Gesneriaceae endemic to the southeastern Atlantic forests of Brazil. Selbyana. 2010; 30:216-220.

24. Vasconcelos, M., and J. Lombardi.Hummingbirds and their flowers in the campos rupestres of southern Espinhaco range, Brazil. Melopsittacus. 2001. 4:3–30.

25. Wolowski M., Ashman, T. and Freitas, L. Community-wide assessment of pollen limitation in hummingbird-pollinated plants of a tropical montane rain forest. Annals of Botany. 2013. 112:903-910.

26. Chautems, A. Révision taxonomique et possibilités d'hybridisations de *Nematanthus* Schrader (Gesneriaceae), genre endémique de la forêt côtière brésilienne. Dissertationes Botanicae. 1988. 112:1-226.

27. Perret, M., A. Chautems, A. O. De Araujo, and N. Salamin. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. Botanical Journal of the Linnean Society. 2013; 171:61–79.

28. Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 2012; 29:1969–1973.

29. Ronquist, F., M. Teslenko, P. van der Mark, D. Ayres, A. Darling, S. Hohna, B. Larget, L. Liu, M. A. Suchard, and J. Huelsenbeck. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology. 2012; 61:539–542.

30. Paradis E., Claude J. & Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004. 20: 289-290.

31. Beaulieu, J. M., B. C. O'Meara, and M. J. Donoghue. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. Systematic Biology. 2013; 62:725–37.

32. Dray, S., and A. B. Dufour. The ade4 package: implementing the duality diagram for ecologists. J. Journal of Statistical Software. 2007; 22:1–20.

33. Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology. 2005; 25:1965–1978.

34. Cooper, N., W. Jetz, and R. P. Freckleton. Phylogenetic comparative approaches for studying niche conservatism. Journal of Evolutionary Biology. 2010; 23:2529–39.

35. Doledec, S., D. Chessel, and C. Gimaret-Carpentier. Niche separation in community analysis: A new method. Ecology. 2000; 81:2914–2927.

36. Harmon Luke J, Jason T Weir, Chad D Brock, Richard E Glor, and Wendell Challenger. GEIGER: investigating evolutionary radiations. Bioinformatics. 2008 24:129-131.

37. Blomberg, S. P., T. Garland, and A. R. Ives. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution. 2003; 57:717–45.

38. Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecoloy and Evolution. 2012; 3:217–223.

39. Revell, L. J. Two new graphical methods for mapping trait evolution on phylogenies. Methods Ecoloy and Evolution. 2013; 4:754–759.

40. Ackerly, D. Conservatism and diversification of plant functional traits : Evolutionary rates versus phylogenetic signal. Proceedings of the National Academy of Sciences of the United States of America. 2009. 106:19699–19706.

41. Hansen, T. F. Stabilizing selection and the comparative analysis of adaptation. Evolution. 1997; 51:1341–1351.

42. Corlay, S. Properties of the Ornstein-Uhlenbeck bridge. 2014; ArXiv**:**1310.5617v4

43. Brown, J. H., and A. Kodric-brown. Convergence, Competition, and Mimicry in a Temperate Community of hummingbird-pollinated flowers. Ecology. 1979; 60:1022–1035.

44. Clark, J. L., and E. A. Zimmer. A preliminary phylogeny of *Alloplectus* (Gesneriaceae): Implications for the evolution of floral resupination. Systematic Biology. 2003; 28:365–375.

45. Clark, J. L. Systematics of *Glossoloma* (Gesneriaceae). Systematic Botany Monographs. 2009; 88:1–128.

46. Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson. Pollination syndromes and floral specialization. Annual Review of Ecology, Evolution, and Systematics. 2004; 35:375–403.

47. Castellanos, M. C., P. Wilson, and J. D. Thomson. “Anti-bee” and “pro-bird” changes during the evolution of hummingbird pollination in Penstemon flowers. Journal of Evolutionary Biology. 2004; 17:876–85.

48. Clark, J.L., Clavijo, L. and Muchhala, N. Convergence of anti-bee pollination mechanisms in the Neotropical plant genus *Drymonia* (Gesneriaceae). Evol. Ecol. 2015. 29:355-377.

49. Muchhala N. and Potts, M. Character displacement among bat-pollinated flowers of the genus *Burmeistera:* analysis of mechanism, process and pattern. Proc. R. Soc. B. 2007. 274: 2731-2737.

50. Armbruster W.S. and Muchhala Nathan. Associations between floral specialization and species diversity: cause, effect, or correlation?. Evolutionary Ecology. 2009. 23:159-179.

51. Muchhala, N., Johnsen, S and Smith, S.D. Pollniation shapes flower color variation in Andean Solanaceae. -Evolution. 2014. 68 (8): 2275-2286.

52. Jakob, S. S., C. Heibl, D. Rödder, and F. R. Blattner. Population demography influences climatic niche evolution: evidence from diploid American *Hordeum* species (Poaceae). Molecular Ecology. 2010; 19:1423–38.

53. Litsios, G., R. O. Wüest, A. Kostikova, F. Forest, C. Lexer, H. P. Linder, P. B. Pearman, N. E. Zimmermann, and N. Salamin. Effects of a fire response trait on diversification in replicated radiations. Evolution. 2013; 68:453–65.

54. Kostikova, A., G. Litsios, N. Salamin, and P. B. Pearman. Linking life-history traits, ecology, and niche breadth evolution in North American eriogonoids (Polygonaceae). The American Naturalist. 2013; 182:760–74.

55. Oliveira-Filho, A. T., and M. A. L. Fontes. Patterns of floristic differentiation among Atlantic forests in southeastern Brazil and the influence of climate. Biotropica. 2000; 32:793–810.

56. Webb, T. J., F. I. Woodward, L. Hannah, and K. J. Gaston. Forest cover-rainfall relationships in a biodiversity hotspot : the Atlantic forest of Brazil. Ecological Applications. 2005; 15:1968–1983.

57. Perret, M., Chautems, A. and Spichiger, R. Dispersal-vicariance analyses in the Tribe Sinningieae (Gesneriaceae): A clue to understanding biogeographical history of the Brazilian Atlantic forest. 2006. Ann. Missouri Bot. Gard. 93:340-358.

58. Perret, M., A. Chautems, R. Spichiger, T. G. Barraclough, and V. Savolainen. The geographical pattern of speciation and floral diversification in the neotropics: the tribe Sinningieae (Gesneriaceae) as a case study. Evolution. 2007; 61:1641–60.

Chapter 2

Hummingbird pollination enhanced the diversification of Gesneriaceae in the Neotropics

This chapter is in review: **Serrano-Serrano ML,** Rolland, J., Clark, JL, Salamin, N, Perret M.

Hummingbird pollination enhanced the diversification of Gesneriaceae in the Neotropics

Authors: Martha Liliana Serrano-Serrano[1,2], Jonathan Rolland[1,2], John L. Clark[3], Nicolas Salamin[1,2*] and Mathieu Perret[4*]

[1]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland, MLSS: martha.serrano@unil.ch, JR: jonathan.rolland@unil.ch, NS: nicolas.salamin@unil.ch.
[2]Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland
[3]Department of Biological Sciences, Box 870345, The University of Alabama, Tuscaloosa, Alabama 35487-0345, U. S. A. JLC: jlc@ua.edu
[4]Conservatoire et Jardin botaniques de la Ville de Genève and Laboratory of Plant Systematics and Biodiversity, University of Geneva, Chemin de l'Impératrice 1, 1292 Chambésy, Geneva, Switzerland, MP: mathieu.perret@ville-ge.ch

* Both authors share the last-authorship

Corresponding author: Mathieu Perret; Conservatoire et Jardin Botanique de la Ville de Genève and Laboratory of Plant Systematics and Biodiversity, University of Geneva, Chemin de l'Impératrice, 1, 1292 Chambésy, Geneva, Switzerland. Email: mathieu.perret@ville-ge.ch. Tél. (+41) 22 418 51 95, Fax (+41) 22 418 51 01.

Summary

* The effects of specific functional groups of pollinators in the diversification of angiosperms are still not clearly known. We investigated whether the pollination shifts or the specific association with hummingbirds affected the diversification of Gesnerioideae in the Neotropics.

* We reconstructed a phylogeny of 583 Gesnerioideae and detected diversification shifts through time, inferred the timing and amount of transitions between pollinator functional groups, and tested the association between hummingbird pollination and speciation and extinction rates in Gesnerioideae.

* We recorded an average of 31.5 transitions to hummingbird pollination and 76.5 reversions to insect pollination. Diversification rates of the group increased through time since 25 Mya, coinciding with the evolution of hummingbird-like flowers and the arrival of hummingbirds in South America. We showed that plants pollinated by hummingbirds have a two-fold higher speciation rate compared to plants pollinated by insects, and that transitions between functional groups of pollinators have a little impact on the process.

* We demonstrate that floral specialization on hummingbirds for pollination has triggered rapid diversification in the Gesnerioideae. Biotic drivers of plant diversification in the Neotropics could be more related to this specific type of pollinator (hummingbirds), than to shifts between different functional groups of pollinators.

Keywords

Co-evolution, comparative methods, floral traits, ornithophyly, pollinator shifts, stochastic mapping.

Introduction

The current species richness of a group of organisms results from the diversification process occurring throughout its evolution. In plants, a variety of intrinsic and extrinsic factors affect the diversification process. The relative importance of these factors has sparked debate and recent studies suggest that diversification is mainly driven by diversity-dependent processes (Rabosky, 2009; Ricklefs, 2007). The availability of geographical areas is important in diversity-dependent diversification as it can alter the ecological limits of a lineage (Vamosi & Vamosi, 2011; Vamosi & Vamosi, 2010). Those limits can however be modified under new climatic conditions (Fiz-Palacios *et al.*, 2011), the colonization of new geographical areas (such as a mountain uplift or island origination; Hughes & Eastwood, 2006; Hughes & Atchison, 2015), or the evolution of particular traits that create new possibilities for species diversification (called "key innovations"; Vamosi & Vamosi, 2010; Litsios *et al.*, 2013, Silvestro et al., 2014). In angiosperms, traits such as biotic pollination, floral symmetry and nectar spurs, which are all related to specialized pollination and the ability to generate reproductive isolation, have been proposed as key innovations due to their positive effects on diversification (Hodges and Arnold 1995; Dodd *et al.*, 1999; Sargent 2004). This support the general idea that specialized biotic pollination is a key factor in the diversification of angiosperms (Stebbins 1974). However, the mechanisms that led to the apparent association between pollination and species richness are still controversial (Armbruster & Muchhala, 2009).

One hypothesis is that diversification in angiosperms has been triggered by the effect of pollinators specialisation on reproductive isolation. Spatial and temporal differences in the availability of the most effective pollinator across the species range could produce pollinator shifts, floral divergence, reproductive isolation, and, ultimately, speciation in plants (reviewed in Kay and Sargent 2009, van der Niet et al. 2014). Evidence for pollinator-shift effects in plant speciation have been found for *Costus* (Kay et al., 2005), *Gladiolus* (Valente *et al.*, 2012), and *Lapeirousia* (Forest *et al.*, 2013), and, a review of available species-level phylogenies estimated that around 25% of the divergence events could be associated with pollinator shifts in angiosperms (Van der Niet and Johnson, 2012). Although these results suggest that frequent pollination shifts have played a role in driving angiosperm diversification, a large proportion of the speciation events could still occurr within specific pollination systems. Indeed, an alternative hypothesis proposes that the diversification rates in angiosperms increases with specialisation on certain guilds of pollinators, rather than with pollinator shifts per se (Valente et al., 2012). For example, vertebrate pollination, and in particular pollination mediated by hummingbirds, is associated with plant species richness in various clades (Schmidt-Lebuhn *et al.*, 2007; Givnish *et al*. 2014; Lagomarsino *et al*., 2016, Roalson and Roberts 2016). This functional group of pollinators could thus be a significant driver of diversification in the Neotropics. These evidences show that both pollinator shifts and transitions within a particular functional group of pollinators can influence plant diversification in angiosperms. However, the relative contribution of these two processes has been rarely tested.

The aim of this study is to evaluate the mode of evolution of functional groups of pollination and their impact on plant diversification in the Neotropical region by focusing on the subfamily Gesnerioideae. This clade of herbaceous plants, shrubs or more rarely small trees contains 75 genera and over 1200 species found exclusively in the Neotropics, with the exception of few Southwest Pacific taxa in the tribe Coronanthereae (Woo *et al.*, 2011; Weber et al., 2013). Based on molecular dating and biogeographical reconstructions, Perret *et al.*,(2013) estimated that Gesnerioideae started its diversification during the early Oligocene, with a rapid range expansion into most Neotropical regions including tropical Andes, Brazilian Atlantic forest (BAF), Cerrado, Central America and the West Indies. The species in this subfamily exhibit a large diversity of floral morphology associated with repeated adaptations to different pollinators such as hummingbird, bees, and bats (SanMartin-Gajardo & Sazima, 2004, 2005a, 2005b; Perret *et al.*, 2007; Martén-Rodríguez *et al.*, 2009, 2015, Clark et al., 2015). Therefore, this clade is particularly interesting to test through which mode and tempo plant-pollinator interactions have evolved and how they influenced species diversification (Roalson and Roberts 2016).

Here, we reconstructed one of the largest species-level phylogeny for a group of Neotropical plants based on four DNA loci and a wide sampling of Gesnerioideae species to test for temporal variations and trait-dependent rates of diversification at a continental scale. Specifically, we assessed whether evolution to hummingbird pollination has contributed to the increased Gesnerioideae diversity in Neotropics and whether it coincides with hummingbird diversification in South America (McGuire *et al.*, 2014). We tested in particular using evolutionary modelling if diversification rates were associated with recurrent shifts of pollinators or if the observed species richness was driven specifically by hummingbird mediated pollination. Addressing these questions in such a large and diversified group of plant will contribute to a better understanding of how ecological factors shape current patterns of species richness in the Neotropics (Givnish *et al.*, 2014).

Material and Methods

Taxonomic sampling and DNA sequencing

Our taxonomic sampling consisted of 583 species representing all the 75 recognized genera in Gesnerioideae and about 50% of the species in the subfamily (Weber *et al.*, 2013). Numbers of species sampled in each tribe and subtribe are as follow: 64 spp. of Beslerieae (out of 220 spp), 284 Columneinae (out of 477), 4 Coronanthereae (out of 20), 38 Gesneriinae (out of 76), 88 Gloxiniinae (out of 167), 14 Napeantheae (out of 22), 86 Ligeriinae (out of 86), 4 Sphaerorrhizinae (out of 4), and 1 Titanotricheae (out of 1). Seven outgroups include representatives of the Didymocarpoideae and Sanangoideae subfamilies, as well as *Peltanthera floribunda* and *Jovellana violacea* (Calceolariaceae) that are close relatives of the Gesneriaceae (Perret *et al.*, 2013).

We obtained DNA sequences for one nuclear (*ITS/5.8S*) and three plastid DNA regions (*matK*, *rps16* intron and *trnL-trnF* intron and spacer) following the procedure described in Araujo *et al.*, (2010) and Perret *et al.*, (2013). A total of 475 sequences were amplified from field samples for this study and merged to available Genbank sequences (see Table S1 for specimen voucher and Genbank information). Sequences were aligned using MAFFT (version 7, Katoh, 2013) and all sites were scored for accuracy of the alignment using Guidance (Penn *et al.*, 2010). We removed sites that had Guidance scores lower than 0.75 and those for which more than 90% of the sequences had missing values or gaps. Our final matrix contained 3,813 base pairs (bp). We identified the best substitution model for each DNA region using the Akaike information criterion (AIC) as implemented in the *phymltest* function in R (ape package; Paradis *et al.*, 2004). We also combined the three plastid markers and used the same procedure to estimate the best fitting model of substitution for this extended partition.

Phylogenetic reconstruction

Topology

Relationships between species were reconstructed by Bayesian inference using MrBayes 3.2 (Ronquist *et al.*, 2012). Two DNA data partitions that corresponded to the ITS/5.8S and the combined chloroplast were used with the associated best fitting model identified by AIC (see above). We performed two runs of the Bayesian inference and each run consisted of four chains of 7 x 107 generations. We sampled the chains every 10^3 generations. We chose the length of the burn-in (20 million steps) and determined the convergence of the MCMC by examining trace plots of each parameter in Tracer v.1.4. (Rambaut *et al.*, 2014). A maximum clade credibility tree with posterior support was calculated by combining the two runs of Bayesian inference using Treeannotator v1.7.0 (Drummond *et al.*, 2012).

Dating

Divergence times were estimated using a relaxed clock model with uncorrelated log-normal prior distribution for the rates of substitution and a birth-death prior for the age of each node as implemented in BEAST v1.7.0 (Drummond *et al*., 2012). We used the same substitution models as described above for the nuclear and plastid DNA partitions. We allowed substitution parameters of each partition to be unlinked, but tree and clock parameters were linked. Secondary calibration was performed by imposing priors for the divergence times for the clade containing all Gesneriaceae (including *Sanango racemosum* and members of the Didymocarpoideae family). We used a log-normal distribution based on a previous analysis that used fossil calibrations from Lamiales (Perret *et al*., 2013, mean crown age of 60.52 Mya and a standard deviation of 1.08 Mya). This is the best estimation available to date in the absence of good fossil record for Gesneriaceae. A total of 70 million generations of MCMC were run and we set a burn-in of 20% based on the inspection of the trace of the MCMC chain using Tracer v.1.4. We generated a Common Ancestor (CA) tree, which prevents the presence of negative branch lengths when averaging node heights compared to the use of the maximum clade credibility tree (Heled & Bouckaert, 2013), using Treeannotator v1.7.0 and sampled randomly 500 trees from the posterior distribution of trees for further analyses (see below).

Characterization of pollination syndromes

The predictability of pollination syndromes is largely debated (Waser *et al.*, 1996, Fenster *et al.*, 2004, Ollerton *et al*., 2009). However, a recent meta-analysis supported the concept of pollination syndromes, especially for tropical plants (Rosas-Guerrero *et al.*, 2014), and encouraged the use of floral characters as a proxy for pollination interactions in macro-evolutionary studies (e.g. Lagomarsino *et al.*, 2016, Roalson and Roberts 2016). In Gesnerioideae, several studies combining field observations and multivariate analyses of morphometric data have demonstrated that suites of floral traits could predict specialized pollination by hummingbirds, bees and bats in *Drymonia* (Clark *et al.*, 2015), Gesneriea (Martén-Rodriguez *et al.*, 2009), *Nematanthus* and *Codonanthe* (Serrano-Serrano *et al.*, 2015), and Sinningieae (Perret *et al.*, 2007).

To further test the validity of pollination syndromes, we assessed the correlation between floral traits and functional groups of pollinators among the species of Gesneriaceae with documented pollination systems. An extended bibliographic search was conducted to identify all Gesnerioideae species with published information about their pollinators (Table S2). Flowers of these species were characterized using nine morphological traits reflecting their variation in size, shape and color (Table S2). Trait values were derived from published morphometric datasets, monographic revisions, and our own measurements of flowers collected in the field or in living collections, or from scaled images available on John L. Clark's website (www.gesneriads.ua.edu). Among these traits, the degree of corolla constriction (i.e. tubular vs bell-shaped corolla) and the presence of pouched or urn-shaped corolla have been identified as key traits to discriminate hummingbird from bee and bat pollinated flowers in different groups of Neotropical Gesneriaceae (Perret *et al*., 2007; Martén-Rodríguez *et al*., 2009; Clark *et al*., 2015; Serrano-Serrano *et al*., 2015). Experimental results have also demonstrated the role of flower constriction and anther exsertion in improving the morphological fit between hummingbirds and flowers and/or in deterring less efficient pollinators such as bees (i.e. anti-bee traits; Castellanos *et al*., 2004). We used a discriminant analyses to maximize the differences in each trait between functional groups of pollinators (i.e. hummingbirds, bats, insects, and generalists), and to estimate their predictability for the identification of pollination/shape associations. We used the *lda* and *predict* functions from MASS R package (Venables and Ripley, 2002). The most discriminant floral traits were then used to predict the functional groups of pollinators for the species included in the phylogeny that lack direct observation of pollinators.

In all subsequent analyses requiring binary states (see below), the bat pollinated species were merged into the hummingbird-pollination syndrome category (8 out of 590 species). We based this choice on the fact that i)

hummingbirds and nectarivory bats are both vertebrates with hovering ability, ii) certain bat pollinated species are generalist (pollinated also by hummingbirds during late afternoon and at dawn; Martén-Rodríguez *et al*., 2009), and iii) according to a three-state stochastic mapping analysis most of the bat pollinated species in Gesnerioideae evolved recently from hummingbird pollinated species (see section below, and Fleming *et al*., 2009).

Evolution of pollination syndromes

The study of trait evolution has largely improved by considering evolutionary time into the modeling of a trait change (Huelsenbeck *et al.,* 2002), and, recently, by including the species diversification process itself (Binary State Speciation and Extinction [BiSSE] models; Maddison *et al.,* 2007). Here, we incorporate most of these improvements by jointly modeling the evolution of pollination syndromes and trait-dependent diversification rates (binary-state trait). For this, we used estimates of transition rates between hummingbird and insect pollination syndromes from the BiSSE model that decomposes the evolutionary process into state-specific speciation and extinction rates and two transition rates. We reconstructed the ancestral states using the asr.marginal function from the diversitree R package (FitzJohn 2012), which is only available for the specific BiSSE model, and not the other extensions of this model (as the ClaSSE model used in the diversification analysis). We mapped changes in pollination syndromes across the Gesnerioideae phylogenetic tree by incorporating the BiSSE estimates of ancestral states into the stochastic mapping (modifying the *simmap* function in phytools R package, Revell, 2012), and ran 200 reconstructions on independent trees. For each stochastic mapping we divided branch lengths into time bins of 1 My and recorded the number of transitions from and to hummingbird pollination syndrome in each bin. We reported the time bin at which 95% of the stochastic mappings have at least one transition event as the onset time for each type of transition. We performed an additional three-state stochastic mapping without considering trait-dependent diversification to explore the evolution among hummingbird, bat and insect pollination syndromes (see Methods S1).

Diversification analysis

Temporal shifts in diversification

We tested whether diversification rates were constant or varied through time using the R package *TreePar* (Stadler, 2011). A sample of 500 Gesnerioideae trees from the posterior distribution estimated by BEAST was used to fit models accounting for shifts in diversification. Models accounted from zero to five shifts, in time bins of 1 My, while accounting for incomplete taxon sampling. Models with *n* and *n+1* shifts were compared using likelihood ratio test until additional shifts did not improve the model fit.

Trait-dependent diversification

We adopted again a trait dependent birth-death model to assess correlations between evolution of pollination syndromes in Gesnerioideae and changes in speciation and extinction rates. We used an extension of the BiSSE model, named ClaSSE (Cladogenetic State change Speciation and Extinction (Goldberg & Igić 2012), and implemented in the R package *diversitree* (FitzJohn 2012) to test for a differential effect of pollination types on diversification rates of Gesneriodeae. This model allows us to infer how diversification rates change in response to shifts in pollination syndromes (switches between insect and hummingbird syndromes) or to interactions within the same pollination syndrome. A binary trait was used to represent the pollination syndromes (insect as state 0; hummingbird as state 1). We estimated six different speciation rates (two to model speciation within pollination syndromes: λ000, λ111; two to model speciation associated with a switch from insect to hummingbird pollination syndrome: λ001, λ011; two to model speciation associated with a switch from hummingbird to insect pollination syndrome: λ100, λ101). We also included two state-

specific extinction rates (μ0, μ1) and two transition rates (q01, q10). We estimated the posterior distributions of each parameter of the ClasSSE model in a Bayesian framework (Silvestro *et al.* 2014) using an exponential prior distribution on the speciation, extinction and transition rates. Defining appropriate priors can be difficult and subjective. We therefore treated the rate of the exponential prior as an unknown variable with a gamma hyper-prior Γ [2, 2] and estimated it from the data sampling it from its conjugate distribution, as described in Silvestro *et al*. (2016). The hierarchical approach used here has the advantage of reducing the risks of over-parameterization through Bayesian shrinkage (Gelman *et al.* 2014) and we compared the posterior distribution of the rate parameters (mean and 95% HPDs) to assess whether they were significantly different from one another. Posterior parameter estimates were obtained through MCMC (2.5×10^4 generations) using 100 trees randomly drawn from the BEAST posterior sample to account for phylogenetic uncertainty. All runs were implemented in R (version 3.2.5) using a script developed by D. Silvestro and M.L. Serrano-Serrano (available at https://github.com/dsilvestro/mcmc-diversitree). Posterior samples were summarized using Tracer (v.1.6, Rambaut *et al.* 2014).

Methods associating traits and diversification (such as ClaSSE) should be taken with caution because high type II error can occur with small phylogenetic trees (<300 species), traits that are highly biased toward one of the states or a low number of statistically independent origination of one character state (Davis *et al.*, 2013; Maddison and FitzJohn, 2015). Our dataset is however robust to these violations because our phylogenetic tree contains almost 600 species, the ratio between character states is 0.68 and the subfamily displays a large number of independent shifts of pollination syndrome. The categorization of the pollination syndrome can be questioned (see "characterization of pollination syndromes" section) and we also tested whether potential misclassification of 1%, 10%, 15%, 20%, and 25% of each pollination syndrome could affect our results. We sampled randomly these percentages of species and changed their pollination syndrome (a hummingbird-pollinated species becoming an insect, and inversely) to account for possible misidentifications in our predictions. We further tested the influence of large clades with a single pollinator type on the diversification results by running the analyses while excluding the *Columnea* genus that contains 86 hummingbird pollination syndrome species. Finally, methods like ClaSSE can also be biased toward detecting false positive relationships between trait and diversification (i.e. type I error; Rabosky and Goldberg, 2015). We evaluated this possible methodological issue by simulating 100 datasets under the null hypothesis that the trait was not associated with diversification, using the *rayDISC* function from the R package corHMM (Beaulieu *et al.*, 2013). We then fitted the ClaSSE model as described above on these 100 simulated datasets and compared the parameter estimation and their posterior distributions obtained under the null hypothesis and our observed data set.

Results

Phylogenetic analysis

The best models of molecular evolution were GTR+Γ and GTR+Γ+I for the nuclear and chloroplast DNA partition, respectively (Fig. S3). The MrBayes and BEAST analyses resulted in congruent topologies. Our phylogenetic reconstruction (Fig. 1 and trees deposited in Treebase) constitutes one of the largest species-level phylogenetic analysis for Neotropical plants. The topology corroborates the formal classification proposed by Weber *et al*. (2013), namely that Gesnerioideae consists in five tribes and 12 subtribes (posterior probabilities >0.99; Fig. 1). Relationships among tribes agreed with Perret *et al*. (2013) with a high support (posterior probabilities >0.99), except that Titanotricheae, Napeantheae and Beslerieae (composed of the genera *Besleria, Gasteranthus, Reldia, Cremosperma, Shuaria, Anetanthus* and *Tylopsacas*) formed a clade (PP =0.508 in the BEAST MCC) sister to the rest of the Gesnerioideae. The tribe Coronanthereae was sister to the Gesnerieae in agreement with prior results (Woo *et al*., 2011). Five highly supported clades were resolved in the Gesnerieae, corresponding to the subtribes Gesneriinae, Gloxiniinae, Columneinae,

Sphaerorrhizinae and Ligeriinae. Generic and infrageneric relationships largely agree with previous phylogenetic results obtained for these lineages (Perret *et al*., 2003; Roalson *et al*., 2008; Araujo *et al*., 2010; Clark *et al*., 2010, 2011, 2012; Serrano-Serrano *et al*., 2015, Mora and Clark, 2016, Araujo *et al.*, In press). Out of 74 genera of Gesnerioideae, 8 appeared non-monophyletic and are still in need of further taxonomical revision (*Achimenes*, *Diastema*, *Gesneria*, *Mandirola*, *Paliavana*, *Phinaea*, *Sinningia*, and *Vanhouttea*).

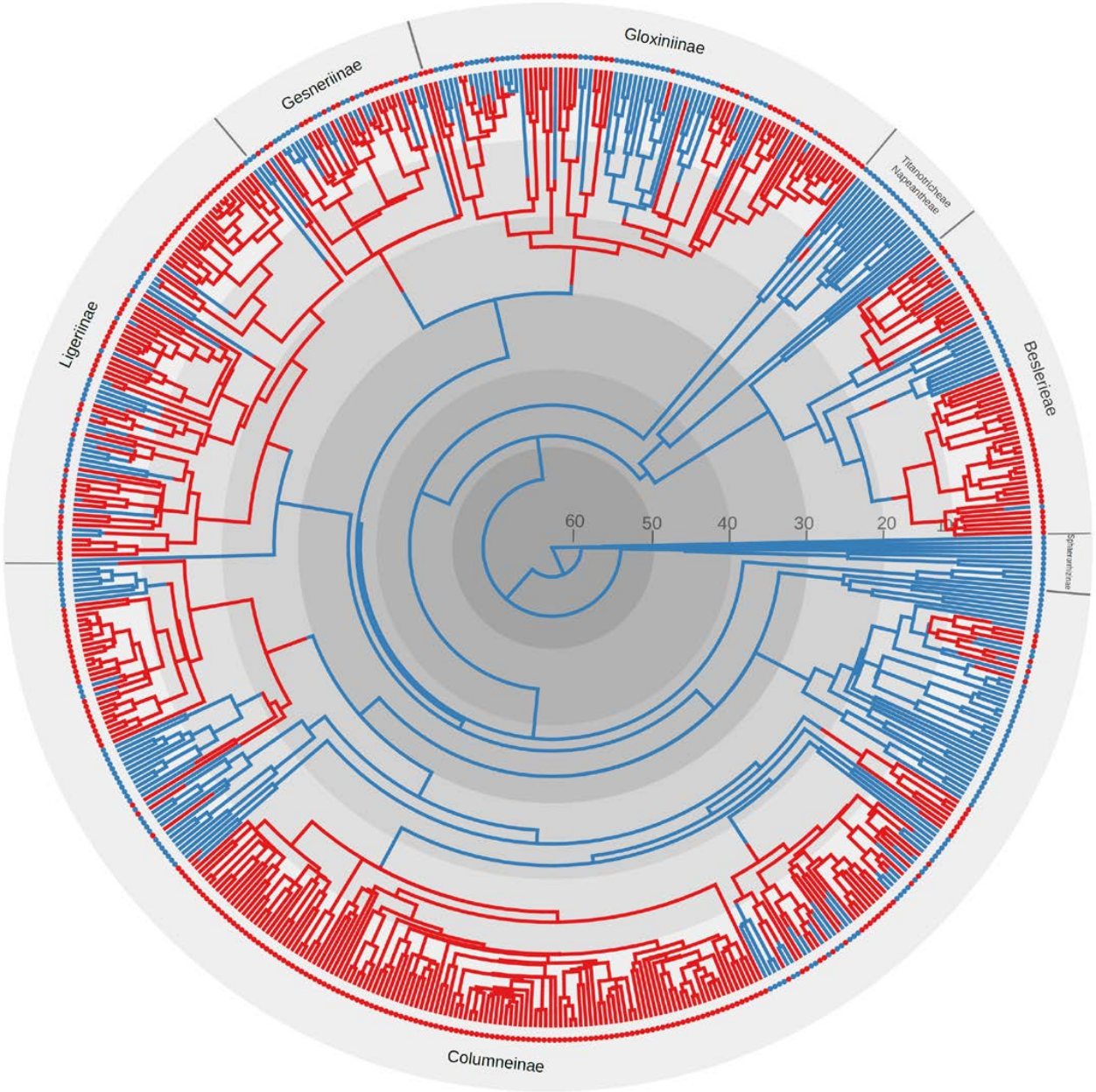


Figure 1. Bayesian Common Ancestor (see in main text) phylogenetic reconstruction showing one stochastic mapping of pollination syndromes. Gray boxes correspond to taxonomic tribes. Colors on branches correspond to pollination syndromes: blue = insect and red = hummingbird. Gray concentric circles have 10 Mya span.

Characterization and evolution of pollination syndromes

Overall, the 118 species with documented pollination systems were recorded from the literature (Table S2). Among them, 82 species were pollinated by hummingbirds, 19 species pollinated by bees, three species pollinated by other insects (butterfly, diptera, and moth), and seven species pollinated by bats (Table S2). Seven other species are pollinated by a mix of nocturnal and diurnal visitors (e.g. hummingbird, bat, and moth). These generalist species of Gesneriaceae have been so far only recorded on the Caribbean islands in pollinator-depauperate environments (Marten-Rodriguez & Fenster, 2010; Marten-Rodriguez *et al.,* 2015) The discriminant analyses explained a large proportion of the floral trait variability (axes 1 and 2 with 98.69% of variance). The predictability of each group of functional pollinators was high (hummingbirds= 0.974, insects=0.954, bats=1.00, generalists=0.66), and their separation in the morphological space was clear (Fig. 2). Only five species have a group predictability lower than 0.8, these are one bee pollinated: *S. villosa*, three generalists: *G. viridiflora, R. leucomallon, R. vernicosum*, and one hummingbird pollinated: *P. sericiflora*, a species with flower morphology related to the bat syndrome but effectively pollinated by hummingbirds (SanMartin-Gajardo and Sazima 2005a). These cases indicate that flower morphology can be sometimes misleading in identifying functional groups of pollinators. Misidentifications are however rare and occurred in less than 5% of the documented species. To evaluate how this bias could affect our analyses we thoroughly tested the robustness of our results to pollinator misidentification (see below). The standardized coefficients of each trait determine the contribution of the respective trait to the discriminant function between the groups. Based on these values (Table S3), we selected tube shape and lobe symmetry (Fig. S1, S2) as a proxy to estimate the number of species pollinated by hummingbird in the phylogeny that have not been studied in the field. Using this approach, and the information listed in Table S2, we inferred 351 species with hummingbird pollination syndrome, 8 species pollinated by bats, and 231 species with insect pollination syndrome among the 590 taxa included in our phylogenetic tree (Table S1).

The BiSSE estimates of transitions rates between pollination syndromes indicated a median rate from insect to hummingbird pollination syndrome of 0.009, and from hummingbird to insect pollination syndrome of 0.044. Our stochastic mapping showed that pollination syndromes evolved on average from insect to hummingbird 31.50 (± 10.07) times. Transitions to hummingbird-pollination syndromes first occurred around 18.5 Mya and then increased in frequency over time (Fig. 3a). These transitions were reconstructed at the crown of major clades of Gesnerioideae, such as *Besleria*, Ligeriineae (*Dircaea*) and *Columnea* (Fig. 1). Reversions from hummingbird to insect pollination were highly frequent (on average 76.50 ± 18.06 times). These reversions to insect pollination started around 12.5 Mya and were mainly reconstructed on terminal branches or within clades including few species (Fig. 1 and 3b). Our three-state reconstruction treating bat and hummingbird syndromes separately showed that transitions to bat pollinated flowers (all observed in the field) have occurred at least seven times, since around 9.5 Mya, and mainly from hummingbird-adapted flowers (Fig. S8, S9).

Diversification analysis

Our analyses of temporal shifts during the diversification of Gesnerioideae detected a single shift in diversification rate (p-value < 0.001, Fig. 4) that most probably occurred around 18.5 Mya (95% confidence interval = 5.0 - 25.5 Mya). The mean net diversification rates were 0.067, and 0.177 Mya^{-1}, for the periods before and after the shift respectively.

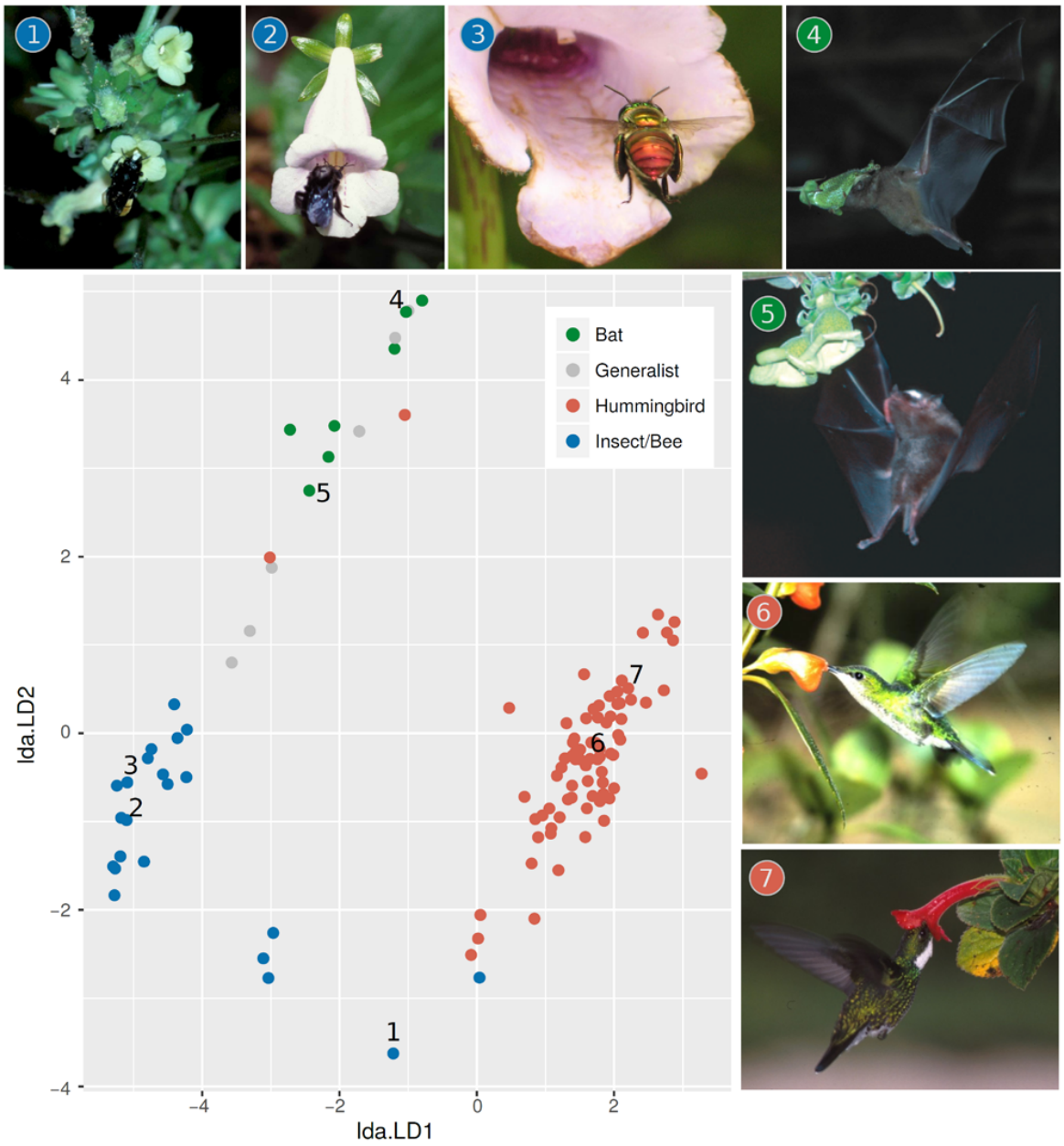


Figure 2. Discriminant analysis conducted on nine floral traits for 118 available species (Table S2). Images 1 to 7 are examples of pollination types in the family. Photo information: 1. *Eufriesea surinamensis* visiting *Sinningia villosa* (by Ivonne SanMartin-Gajardo in SanMartin-Gajardo & Sazima, 2004); 2. *Bombus morio* visiting *Sinningia eumorpha* (by Ivonne SanMartin-Gajardo in SanMartin-Gajardo & Sazima, 2004); 3. *Euglossa* visiting *Gloxinia perennis* (by Anton weber in Witschnig, *et al*. 2008); 4. *Glossophaga soricina* on flowers of *Sinningia brasiliensis* (by Ivonne SanMartin-Gajardo in SanMartin-Gajardo & Sazima, 2005); 5. *Anoura caudifer* visiting *Paliavana prasinata* (by Ivonne SanMartin-Gajardo in SanMartin-Gajardo & Sazima, 2005); 6. *Thalurania glaucopis* visiting *Nematanthus fornix* (by Leandro Freitas in Wolowski *et al.*, 2013); 7. *Leucochloris albicollis* visiting *Vanhouttea hilariana* (by Ivonne SanMartin-Gajardo in SanMartin-Gajardo & Sazima, 2005).

The analyses of trait-dependent diversification based on the ClaSSE model suggested that speciation rates within pollination syndromes were higher than those associated with shifts between them (Figure 5). Further, species within the hummingbird pollination syndrome have at least a two-fold higher rate (mean λ111 = 0.252 Myr^{-1}, 95% HPD = 0.193 – 0.314) than species within the insect pollination syndrome (mean λ000 = 0.102 Myr^{-1}, 95% HPD = 0.071 – 0.133). All speciation rates associated with a shift in pollination syndrome, regardless the direction of the shifts, are lower or close to 0.01 and thus an order of magnitude lower than the rates within pollination syndrome (λ001=0.006, λ011=0.003, λ101=0.004, λ100=0.010; Figure 5). Posterior distributions of extinction rate showed a higher extinction rate for hummingbird-pollination syndrome species (mean μ0 = 0.015, 95% HPD = 0.000 – 0.041, and μ1 =0.027, 95% HPD= 0.000 – 0.077). Transition rates between pollination syndrome states supported a higher rate of reversals to insect-pollination syndrome (mean q01 = 0.006, HPD = 0.000 – 0.009, and mean q10= 0.023, HPD = 0.000 – 0.040), and were of similar magnitude than the rates estimated by BiSSE (Table S4).

Table 1. Mean and 95% HPD for the parameters of the ClaSSE model. λ= speciation rate, μ = extinction rate, and q = transition rates between states, 0 = insect, 1= hummingbird.

Parameters	Mean	95% HPD
Speciation rate within insect pollination syndrome (λ_{000})	0.102	0.071 – 0.133
Speciation rate with pollination syndrome shift (λ_{001})	0.006	0.000 – 0.013
Speciation rate with pollination syndrome shift (λ_{011})	0.003	0.000 – 0.006
Speciation rate with pollination syndrome shift (λ_{100})	0.010	0.000 – 0.021
Speciation rate with pollination syndrome shift (λ_{101})	0.004	0.000 – 0.016
Speciation rate within hummingbird pollination syndrome (λ_{111})	0.252	0.193 – 0.314
Extinction rate insect pollination syndrome (μ_0)	0.015	0.000 – 0.041
Extinction rate hummingbird pollination syndrome (μ_1)	0.027	0.000 – 0.077
Transition rate from insect to hummingbird pollination syndrome (q_{01})	0.006	0.000 – 0.009
Transition rate from hummingbird to insect pollination syndrome (q_{10})	0.023	0.000 – 0.0400

We found that diversification results are robust to the misidentification of functional groups of pollinators at the tips of phylogenetic tree. First, the difference in rates of speciation between the two pollination syndromes (λ000 and λ111) is persistent if we remove the species-rich genus *Columnea*, which includes exclusively humming-pollination syndrome species (Fig. S4a). Second, the test for possible misidentification of functional groups of pollinators indicated that our estimation of speciation and extinction rates are extremely robust to up to 10% of equivocal states (for both insect- and hummingbird) and that even 15% of misidentification leads to

qualitatively similar results (Figure S5, S6, and S7). Finally, the simulations of traits, whose evolution is independent from the diversification process, showed that the estimated speciation rates within pollination syndromes (i.e. λ000 and λ111) are equal, under the null hypothesis, as well as the extinction and transition rates (Fig. S4b). The effect of hummingbird pollination syndrome on diversification that we detected is thus not likely due to a particular shape of the phylogeny (as suggested in some cases by Rabosky and Goldberg, 2015) that could lead to a false detection of an association between traits and speciation.

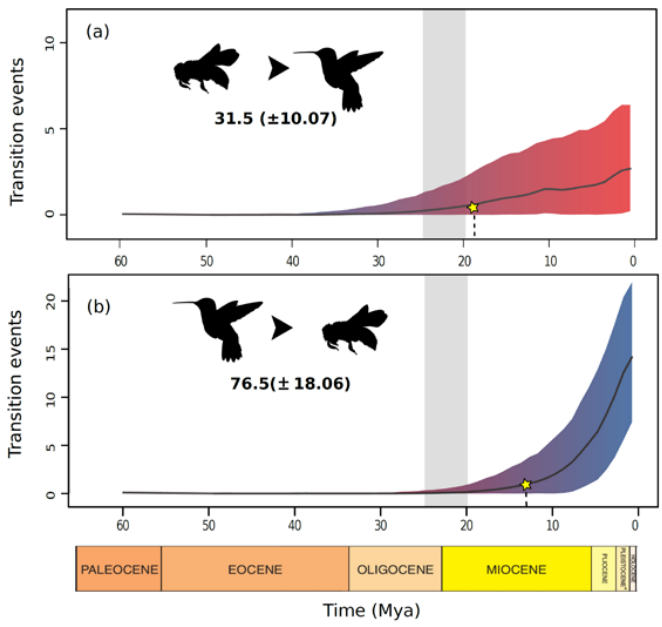


Figure 3. Estimated number of transitions through time for pollination syndromes. Numbers below the pictograms correspond to the mean total number of transitions between the states and the standard deviation. Starts denote the starting point in time where at least one transition is recorded in 95% of the reconstructions. Gray bar is the age of the most-recent common ancestor of extant hummingbirds (20.3–24.7 Mya; McGuire *et al.* 2014).

Discussion

This study showed that hummingbird pollination likely played a role in the diversification dynamics of Gesnerioideae in the Neotropics. Two lines of evidence support this result. First, the diversification of this subfamily increased substantially around 20 Mya. This period corresponds closely to the dispersal of hummingbirds into South America (McGuire *et al.*, 2014) and the first appearance of plant species with hummingbird-pollination syndrome in the Gesnerioideae (Fig. 3). Second, we clearly show that species with hummingbird pollination syndrome have higher rates of speciation compared to species with insect-pollination syndrome. We can thus conclude that the evolution of floral traits associated with hummingbird pollination, and ultimately this biotic interaction in multiple lineages, may have increased the rates of diversification of Gesnerioideae.

Evolution of hummingbird pollination syndrome in gesnerioideae

Our study reveal that Gesnerioideae was originally pollinated by insects and that at least 31 transitions to hummingbirds and bat pollination syndromes occurred during its evolution (Fig. 1 and S9). The repeated evolution of hummingbird pollination syndrome in independent Gesnerioideae lineages centered into different geographical areas, such as the Brazilian Atlantic forest, Andes, Caribbean islands, and Central America (Perret *et al.*, 2013), is indicative of the success of this ecological interaction across the entire Neotropics. We found also frequent state reversals from hummingbird to insect pollination syndrome contradicting the idea that the evolution of hummingbird pollination could act as a dead-end from where reversals to other pollination modes are no longer occurring (Wilson *et al.*, 2007; Tripp & Manos, 2008; van der Niet & Johnson, 2012;

Barrett, 2013). The ability of Gesnerioideae species to change their floral morphologies and pigmentation to such distinctive types in relatively short periods of time (Fig. 2), as well as the reversibility of this system, are striking and motivates the investigation of the genetic mechanisms controlling these transitions (Stuurman *et al.*, 2004, Cronk & Ojeda, 2008; Wessinger and Rausher, 2014).

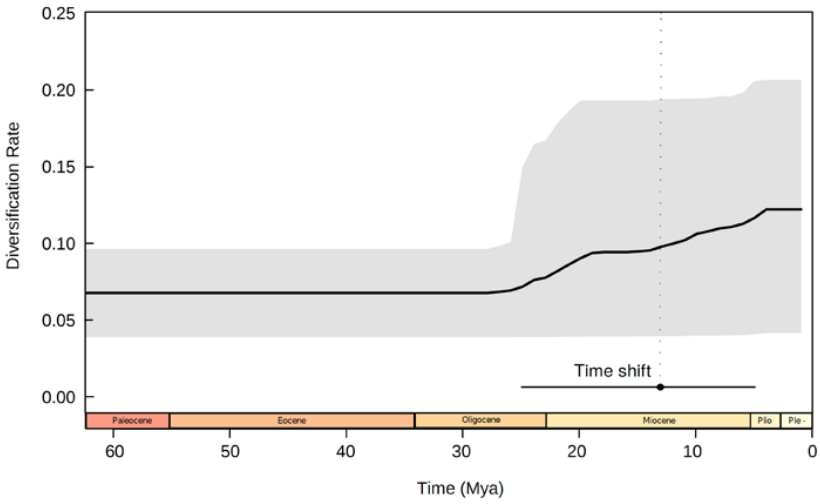


Figure 4. Diversification rate estimates for the time variation analysis (*TreePar*). Gray shadow is the 95% and black line the mean diversification rate across reconstructions. Horizontal black line represent the distribution of time for the estimated shift across trees (5 – 25.05 Mya), black point and vertical dotted line are the mean value for the time shift (13.16 Mya).

Flowers with a morphology corresponding to a hummingbird-pollination syndrome appeared in Gesnerioideae around 18.5 Mya (Fig. 2a). This date is close to the first colonization of hummingbirds in South America and the onset of their diversification on this continent (22.4 Mya; McGuire *et al.*, 2014). This early origin of hummingbird flowers, and the inferred south American origin of the Gesnerioideae (Perret *et al.* 2013), indicate that this plant group could have interacted with the first hummingbirds living in South America, unlike more recent hummingbird-adapted plant lineages (e.g., *Ruellia*, Tripp & McDade, 2013; Bromeliaceae, Givnish et al., 2014; Campanulaceae, Lagomarsino *et al.*, 2016). Our results provide also evidence that plants and hummingbirds have interacted during a longer period of time in tropical South America than in Northern America and temperate South America, two regions that host younger assemblages of hummingbirds and hummingbird-adapted species dated to 6-7 Mya and 16-17 Mya respectively (Abrahamczyk and Renner, 2015).

Transitions between insects and hummingbird flowers were however not clustered in time but occurred with an increased frequency over time since the Miocene (Fig. 3). This pattern parallels the hummingbird species accumulation overtime into the different American biomes, and especially in the Andes (McGuire *et al.*, 2014). Overall, this co-diversification between Gesnerioideae and hummingbirds and the apparent weak specialization of these plants on specific hummingbird species (Sazima *et al.*, 1996, SanMartin-Gajardo and Sazima 2005b) correspond to a process described as diffuse co-evolution (Janzen, 1980; Tripp and McDade, 2013). However, further analysis reconstructing the history of local plant-hummingbird assemblages will be required to better understand the evolutionary dynamics of this mutualism in time and space (e.g., Graham *et al.*, 2010; Abrahamczyk et al. 2015).

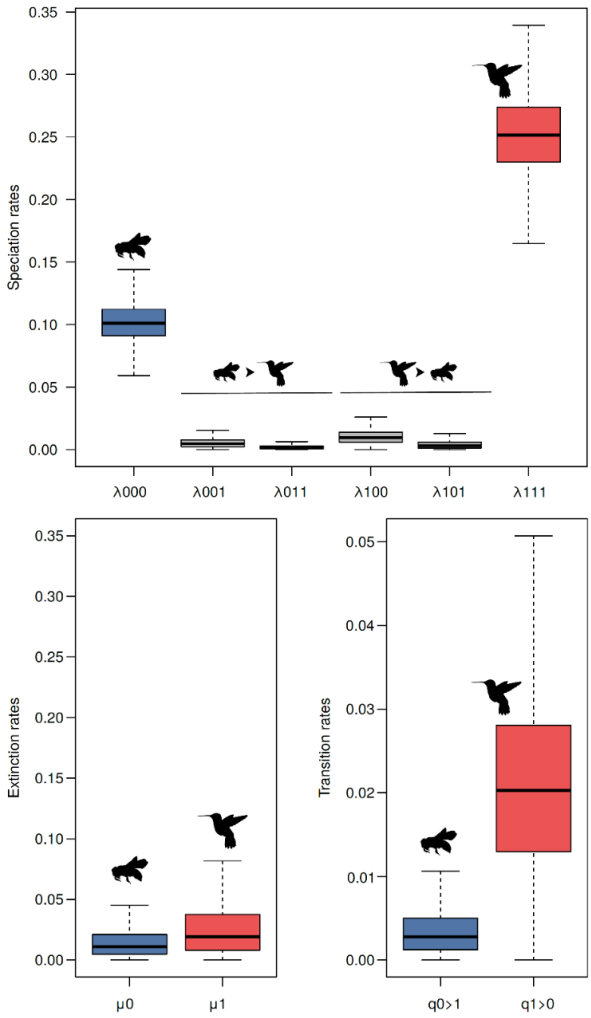


Figure 5. State-dependent speciation (a), extinction (b) and transition (c) rate estimates from ClaSSE model. See parameter description in table 1. Colors and pictograms correspond to the binary pollination syndrome: blue for bee, and red for hummingbird states.

Effect of hummingbird pollination on diversification

Our finding of a two-fold increase in speciation rates for species with a hummingbird-pollination compared with insect-pollination syndrome suggests that floral morphologies associated with hummingbird pollinators may promote mechanisms that lead to the generation of new species. By contrast, speciation rates associated with shifts in pollination syndromes (i.e. between insect and hummingbird pollination syndromes) were between 20 and 80 times lower than those within pollination syndromes (see table 1). These results indicate that the classical pollinator-shift model driving plant speciation is not the major process explaining Gesnerioideae diversification. Instead, species richness in this plant group has been mainly driven by speciation within hummingbird pollinated linages, without involving shifts between functional groups of pollinators. This result agrees with previous results identifying a positive effect of hummingbird pollination syndrome on speciation rates (Givnish *et al.*, 2014, Lagomarsino et al., 2016, Roalson and Roberts, 2016) suggesting a more global effect of this biotic interactions on Neotropical plant diversity.

Why hummingbird pollination promotes plant speciation remains unclear (Schmidt-Lebuhn *et al*., 2007) and several non-exclusive hypotheses exist. First, the evolution of tubular or gullet-like flowers characterizing most hummingbird flowers may have directly accelerate speciation by promoting specialised relationships with the different bill-length categories of hummingbird species and the evolution of rapid prezygotic reproductive barriers (Givnish 2010). Second, flower specialisation and specific pollen placement on hummingbird body may also prevent interspecific pollen transfer between species sharing similar pollinators thereby influencing

the number of species that can co-occur in a same community (Brown and Kodric-Brown 1979; Sargent and Ackerly 2008, Serrano-Serrano et al., 2015; Temeles *et al.,* 2016). It has been suggested that this process could decrease extinction rates (Armbruster & Muchhala, 2009), but also potentially increase the carrying capacity of hummingbird pollinated lineages per unit of area, a factor that can limit the decline of diversification rates over time (Vamosi *et al.,* 2014). Third, hummingbird pollination could also affect the range of pollen dispersal and therefore the connectivity between natural populations (Castellanos et al., 2003). For instance, the South African Gesneriaceae *Streptocarpus primulifolius*, whose pollination is mediated by a fly, displayed a reduced pollen dispersal compared to the related (non-sister species) sunbird pollinated *S. dunnii* (Hughes *et al.,* 2007). Evaluating the differences in geographic range between related species pollinated by hummingbirds could show whether this pollination type contributes to the establishment of geographically isolated or climatically specialized populations (see Perret et al., 2007; Abrahamczyk *et al.* 2014, Schnitzler *et al.,* 2011). Finally, hummingbird pollination is supposed to be more efficient than insect pollination in Neotropical cloud forests at middle to high elevations, because insects are indeed less active in cool, foggy, and wet conditions (Armbruster & Berg, 1994, Cruden, 1972). This capability suggests that hummingbird pollinated species could have more opportunities to persist and speciate in mountain systems compared to insect pollinated lineages. These hypotheses remain so far mostly untested. Further development along these lines will require more complete morphological characterization of the plant species, and plant-pollinator community data, to better understand how biotic interactions have shaped biodiversity and macro-evolutionary patterns in the Neotropical region.

Conclusions

We identified a strong and positive effect of hummingbird-pollination syndrome on the process of species diversification in the subfamily Gesnerioideae. This effect has likely been triggered by the repeated acquisition of hummingbird pollination as soon as this pollination niche became available in South America around 22 Mya. Plants within hummingbird pollination syndrome have increased by twofold the rate of speciation suggesting a positive effect of hummingbird pollination on the establishment of isolation mechanisms. Our findings complement the global understanding of the diversification processes leading to the exceptional diversity of flowering plants in the Neotropics (Antonelli and Sanmartín, 2011; Hughes *et al.,* 2013), and provide new directions towards further testing the role played by plant-pollinator relationships in the build-up of plant diversity.

Acknowledgements

We thank Alain Chautems, Lianka Cairampoma, Andrea O. De Araujo, Valquiria Ferreira Dutra, Mauro Peixoto, Angela Cano, Harri Lorenzi, Gabriel E. Ferreira, Juvenal E. Batista, Mireya Correa, and the staff of the Smithsonian Tropical Research Institute for their contribution to the field work; Régine Niba and Fadil Avdija for their contribution to the laboratory work; Yvonne Menneret, Marina Magnette, Matthieu Grillet, Alexandre Chappuis, Bertrand Guigon and Vincent Goldschmid for the propagation and maintenance of the Gesneriaceae collection at Geneva. Collection permits were granted by the CNPq in Brazil (CMC 038/03) and in the ANAM in Panama (SC/P-43-10). Images of the Figure 2 were kindly provided by Ivonne SanMartin-Gajardo, Anton Weber and Marina Wolowski. We thank Daniele Silvestro for the helpful discussions and support for the analyses. All analyses were run at the High-performance Computing Center (Vital-IT) from the Swiss Institute of Bioinformatics. This study was funded by the Ville de Genève, the Faculty of Biology and Medicine at the University of Lausanne and the Swiss National Science foundation (grant number CRSII3_147630).

Author contribution

M.L.S.S., N.S. and M.P. planned and designed the research. M.P. and J.L.C. conducted fieldwork and gathered data. M.L.S.S., M.P. and N.S. performed the analyses. M.L.S.S., J.R., J.L.C., N.S. and M.P. wrote the manuscript.

References

Abrahamczyk S, Renner SS. 2015. The temporal build-up of hummingbird/ plant mutualisms in North America and temperate South America. *BMC Evolutionary Biology* **15:**104.

Abrahamczyk S, Souto-Vilarós D, McGuire JA, Renner SS. **2015**. Diversity and clade ages of West Indian hummingbirds and the largest plant clades dependent on them: a 5-9 Myr young mutualistic system. *Biological Journal of the Linnean Society* **114**: 848–859.

Abrahamczyk S, Souto-Vilarós D, Renner SS. 2014. Escape from extreme specialization: Passion-flowers, bats and the sword-billed hummingbird. *Proceedings of the Royal Society B* **281**: 20140888.

Antonelli A, Sanmartín I. **2011**. Why are there so many plant species in the Neotropics? *Taxon* **60**: 403–414.

Araújo AO, Castro VC, Perret M. 2010. *Chautemsia calcicola*: A new genus and species of Gloxinieae (Gesneriaceae) from Minas Gerais, Brazil. *Taxon*. **59**: 203–208.

Araújo AO, Chautems A, Cardoso-Gustavson P, Castro Souza V, Perret M. Taxonomic Revision and Phylogenetic Position of the Brazilian Endemic Genus *Sphaerorrhiza* (Sphaerorrhizineae, Gesneriaceae) Including Two New Species. In press in *Systematic Botany.*

Armbruster WS, Berg EE. 1994. Thermal ecology of male euglossine bees (Apidae: Euglossini) in a tropical wet forest: fragrance foraging in relation to operative temperatures. *Biotropica* **26**:50–60.

Armbruster WS, Muchhala N. 2009. Associations between floral specialization and species diversity: cause, effect, or correlation? *Evolutionary Ecology*. **23**: 159–179.

Barrett, SCH. 2013. The evolution of plant reproductive systems: how often are transitions irreversible? *Proceedings of the Royal Society B* **280**:20130913.

Beaulieu JM, O'Meara BC, Donoghue MJ. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Systematic Biology* **62**: 725–37.

Brown JH, Kodric-Brown A. 1979. Convergence, competition, and mimicry in a temperate community of hummingbird-pollinated flowers. *Ecology* **60**: 1022– 1035.

Castellanos MC, Wilson P, Thomson JD. 2003 Pollen transfer by hummingbirds and bumblebees, and the divergence of pollination modes. *Evolution* **57**, 2742– 2752.

Castellanos MC, Wilson P, Thomson JD. **2004**. "Anti-bee" and "pro-bird" changes during the evolution of hummingbird pollination in *Penstemon* flowers. *Journal of Evolutionary Biology* **17**: 876–85.

Clark, JL, Neill DA, Gruhn JA, Weber A, Katan T. 2010. *Shuaria* (Gesneriaceae), an arborescent new genus from the Cordillera del Cóndor and Amazonian Ecuador. *Systematic Botany* **35**(3): 662–674.

Clark, JL, Roalson EH, Pritchard RA, Coleman CL, Teoh V, Matos J. 2011. Independent origin of radial floral symmetry in the Gloxinieae (Gesnerioideae: Gesneriaceae) is supported by the rediscovery of *Phinaea pulchella* in Cuba. *Systematic Botany* **36**: 757–767.

Clark JL, Funke MM, Duffy AM, Smith JF. 2012. Phylogeny of a Neotropical clade in the Gesneriaceae: More tales of convergent evolution. *International Journal of Plant Sciences* **173**: 894–916.

Clark JL, Clavijo L, Muchhala N. 2015. Convergence of anti-bee pollination mechanisms in the Neotropical plant genus *Drymonia* (Gesneriaceae). *Evolutionary Ecology* **29**: 355-377.

Cruden RW. 1972. Pollinators in High-Elevation Ecosystems: Relative Effectiveness of Birds and Bees. *Science*. **176** (4042): 1439-1440

Cronk Q, Ojeda I. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* **59**: 715–727.

Davis MP, Midford PE, Maddison W. **2013**. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology* **13**: 38.

Dodd ME, Silvertown J, Chase MW. **1999**. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* **53**: 732–744.

Drummond AJ, Suchard MA, Xie D, Rambaut A. **2012**. Bayesian Phylogenetics with BEAUti and the

BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.
Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. **2004**. Pollination syndromes and floral specialization. *Annual Review of Ecology, Evolution, and Systematics* **35**: 375–403.
Fiz-Palacios O, Schneider H, Heinrichs J, Savolainen V. 2011. Diversification of land plants: Insights from a family-level phylogenetic analysis. *BMC Evolutionary Biology* **11**: 341.
FitzJohn, R.G. 2012. Diversitree: Comparative Phylogenetic Analyses of Diversification in R. *Methods in Ecology and Evolution*, **3**: 1084-1092.
Fleming TH, Geiselman C, Kress WJ. **2009**. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017–1043.
Forest F, Goldblatt P, Manning JC, Baker D, Colville JF, Devey DS, Jose S, Kaye M, Buerki S. 2013. Pollinator shifts as triggers of speciation in painted petal irises (*Lapeirousia*: Iridaceae). *Annals of Botany* **113**: 357–371.
Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. 2013. Bayesian Data Analysis, Third edition. *Chapman and Hall, CRC Press.*
Givnish TJ. 2010. Ecology of plant speciation. *Taxon* **59**(5):1326–1366
Givnish TJ, Barfuss MHJ, Van EB, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JAC, *et al.* 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution* **71**: 55–78.
Goldberg EE and Igíc B. 2012. Tempo and mode in plant breeding system evolution. *Evolution.* **66** (12): 3701-3709.
Graham CH, Parra JL, Rahbek C, McGuire JA. 2009. Phylogenetic structure in tropical hummingbird communities. *Proceedings of the National Academy of Sciences USA.* **106** (2): 19673-19678.
Heled J, Bouckaert RR. **2013**. Looking for trees in the forest: Summary tree from posterior samples. *BMC Evolutionary Biology* **13**: 221.
Hodges SA, Arnold ML. **1995**. Spurring plant diversification: Are floral nectar spurs a key innovation ? *Proceedings of the Royal Society B* **262**: 343–348.
Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic mapping of morphological characters. *Systematic Biology.* **52** (2): 131-158.
Hughes CE, Eastwood R. **2006**. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences USA* **103**: 10334–10339.
Hughes CE, Pennington RT, Antonelli A. **2013**. Neotropical plant evolution: Assembling the big picture. *Botanical journal of the Linnean Society* **171**: 1–18.
Hughes CE, Atchison, GW. 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytologist* **207:** 275–282.
Hughes M, Möller M, Edwards T, Bellstedt DU, de Villiers M. 2007. The impact of pollination syndrome and habitat on gene flow: a comparative study of two *Streptocarpus* (Gesneriaceae) species. *American Journal of Botany.* **94(10):** 1688-1695.
Janzen DH. 1980. When is it coevolution? *Evolution* **34**:611–612
Katoh, S. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780.
Kay KM, Reeves PA, Olmstead RG, Schemske DW. **2005**. Rapid speciation and the evolution of hummingbird pollination in Neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *American Journal of Botany* **92**: 1899–1910.
Kay KM, Sargent RD. 2009. The role of animal pollination in plant speciation: Integrating ecology, geography, and genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**:637–56.
Lagomarsino LP, Condamine FL, Antonelli A, Mulch A, Davis CC. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist.* doi:10.1111/nph.13920
Litsios G, Wüest RO, Kostikova A, Forest F, Lexer C, Linder HP, Pearman PB, Zimmermann NE,

Salamin N. **2013**. Effects of a fire response trait on diversification in replicated radiations. *Evolution* **68**: 453–65.
Maddison WP, FitzJohn RG. **2015**. The unsolved challenge to phylogenetic correlation test for categorical characters. *Systematic Biology* **64**: 127–136.
Maddison WP, Midford PE, Otto SP. **2007**. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* **56**: 701–710.
Martén-Rodríguez S, Almarales-Castro A, Fenster CB. **2009**. Evaluation of pollination syndromes in Antillean Gesneriaceae: Evidence for bat, hummingbird and generalized flowers. *Journal of Ecology* **97**: 348–359.
Martén-Rodríguez S, Quesada M, Castro A, Lopezaraiza-Mikel M, Fenster CB. 2015. As comparison of reproductive strategies between island and mainland Caribbean Gesneriaceae. *Jorunal of Ecology*. **103:** 1190-1204.
Martén-Rodríguez S, Fenster CB. 2010. Pollen limitation and reproductive assurance in Antillean Gesnerieae: A specialists vs. generalist comparison. *Ecology* **91**: 155–165.
McGuire JA, Witt CC, Remsen JV, Corl A, Rabosky DL, Altshuler DL, Dudley R. **2014**. Molecular phylogenetics and the diversification of hummingbirds. *Current Biology* **24**: 910–916.
Mora MM and Clark JL. 2016. Molecular phylogeny of the Neotropical genus *Paradrymonia* (Gesneriaceae), reexamination of generic concepts and the resurrection of *Trichodrymonia* and *Centrosolenia.* DOI 10.1600/036364416X690561.
Ollerton J, Alarcón R, Waser NM, Price M V., Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. **2009**. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**: 1471–1480.
Paradis E, Claude J, Strimmer K. **2004**. APE: Analyses of Phylogenetics and Evolution in language. *Bioinformatics* **20**: 289–290.
Penn O, Privman E, Ashkenazy H, Landan G, Graur D, Pupko T. **2010**. GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Research* **38**: 23–28.
Perret M, Chautems A, De Araujo AO, Salamin N. **2013**. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. *Botanical Journal of the Linnean Society* **171**: 61–79.
Perret M, Chautems A, Rodolphe S, Kite G, Savolainen V. **2003**. Systematics and evolution of tribe Sinningieae (Gesneriaceae): Evidence from phylogenetic analysis of six plastid DNA regions and nuclear ncpGS. *American Journal of Botany* **90**: 445–460.
Perret M, Chautems A, Spichiger R, Barraclough TG, Savolainen V. **2007**. The geographical pattern of speciation and floral diversification in the Neotropics: The tribe Sinningieae (Gesneriaceae) as a case study. *Evolution* **61**: 1641–60.
Rabosky DL. 2009. Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters* **12**: 735–743.
Rabosky DL, Goldberg EE. **2015**. Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* **64**: 340–355.
Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer
Revell LJ. **2012**. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.
Ricklefs RE. 2007. History and diversity: Explorations at the intersection of ecology and evolution. *American Naturalist* **170**: S56-S70.
Roalson EH, Skog LE, Zimmer EA. 2008. Untangling Gloxinieae (Gesneriaceae). II. Reconstructing biogeographic patterns and estimating divergence times among New World continental and island lineages. *Systematic Botany* **33**:159-175.
Roalson EH and Roberts WR. 2016. Distinct processes drive diversification in different clades of Gesneriaceae. *Systematic Biology. doi: 10.1093/sysbio/syw012*
Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Hohna S, Larget B, Liu L, Suchard MA,

Huelsenbeck J. **2012**. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539–542.

Rosas-Guerrero V, Aguilar R, Marten-Rodriguez S, Ashworth L, Lopezaraiza-Mikel M, Bastida JM, Quesada M. 2014. A quantitative review of pollination syndromes: do floral traits predict effective pollinators? *Ecology Letters*. **17:** 388-400.

SanMartin-Gajardo I, Sazima M. **2004**. Non-Euglossine bees also function as pollinators of *Sinningia* species (Gesneriaceae) in southeastern Brazil. *Plant biology* **6**: 506–12.

Sanmartin-Gajardo I, Sazima M. **2005a**. Chiropterophily in sinningieae (Gesneriaceae): *Sinningia brasiliensis* and *Paliavana prasinata* are bat-pollinated, but *P. sericiflora* is not. Not yet? *Annals of Botany* **95**: 1097–1103.

Sanmartin-Gajardo I. and M. Sazima. 2005b. Espécies de *Vanhouttea* Lem. e *Sinningia* Nees (Gesneriaceae) polinizadas por beija-flores: interações relationadas ao hábitat da planta e ao néctar. *Revista Brasileira de Botânica* **28**:441-450.

Sargent R. 2004. Floral symmetry affects speciation rates in angiosperms. *Proceedings of the Royal Society of London. Series B.* **271**. 603-608.

Sargent RD, Ackerly DD. 2008 Plant–pollinator interactions and the assembly of plant communities. *Trends in Ecology & Evolution*, **23**, 123–130

Sazima I, Buzato S, Sazima M. 1996. An Assemblage of Hummingbird-pollinated Flowers in a Montane Forest in Southeastern Brazil**.** *Botanica Acta* **109:** 149–160.

Schmidt-Lebuhn AN, Kessler M, Isabel H. **2007**. Hummingbirds as drivers of plant speciation? *Trends in Plant Science* **12**: 329–331.

Schnitzler J, Barraclough TG, Boatwright JS, Goldblatt P, Manning JC, Powell MP, Rebelo T, Savolainen V. 2011. Causes of plant diversification in the Cape biodiversity hotspot of South Africa. *Systematic Biology*. **60 (3):** 343-357.

Serrano-Serrano ML, Perret M, Guignard M, Chautems A, Silvestro D, Salamin N. 2015. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evolutionary Biology* **15**: 247.

Silvestro D, Zizka G, Schulte K. **2014**. Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). *Evolution* **68**: 163–75.

Silvestro D, Zizka A, Bacon CD, Cascales-Miñana B, Salamin N, Antonelli A. 2016. Fossil biogeography: a new model to infer dispersal, extinction and sampling from palaeontological data. *Philosophical transactions of the Royal Society B.* DOI: 10.1098/rstb.2015.0225

Stadler T. **2011**. Mammalian phylogeny reveals recent diversification rate shifts. *Proceedings of the National Academy of Sciences of the USA* **108**: 6187–92.

Stebbins , G. L. 1974 . Flowering plants: Evolution above the species level. *Belknap Press of Harvard University Press*, Cambridge, Massachusetts, USA.

Stuurman J, Hoballah ME, Broger L, Moore J, Basten C, Kuhlemeier C. 2004. Dissection of floral pollination syndromes in Petunia. *Genetics* **168:**1585–99

Temeles EJ, Newman JT, Newman JH, Cho SY, Mazzota AR, Kress WJ. 2016. Pollinator competition as a driver of floral divergence: An experimental test. *PLOS ONE*. **11** (1): 1-17.

Tripp EA, Manos PS. 2008. Is floral specialization an evolutionary dead-end? Pollination system transitions in *Ruellia* (Acanthaceae). *Evolution.* **62**: 1712–1737.

Tripp EA, Mcdade LA. **2013**. Time-calibrated phylogenies of hummingbirds and hummingbird-pollinated plants reject a hypothesis of diffuse co-evolution. *Aliso* **31**: 89–103.

Valente LM, Manning JC, Goldblatt P, Vargas P. 2012. Did Pollination Shifts Drive Diversification in Southern African Gladiolus? Evaluating the Model of Pollinator-Driven Speciation. *The American Naturalist.* **180**(1): 83-98

Vamosi JC, Vamosi SM. 2010. Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. *Ecology Letters* **13**: 1270–1279.

Vamosi JC, Vamosi SM. **2011**. Factors influencing diversification in angiosperms: At the crossroads of intrinsic and extrinsic traits. *American Journal of Botany* **98**: 460–71.
Vamosi JC, Armbruster WS, Renner SS. 2014. Evolutionary ecology of specialization: insights from phylogenetic analysis. *Proceedings of the Royal Society B.* **281**. 20142004.
Van der Niet T, Johnson SD. 2012. Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends in Ecology and Evolution* **27**: 353–361.
Van der Niet T, Johnson SD. 2014. Pollinator-driven ecological speciation in plants: new evidence and future perspectives. *Annals of Botany* **113**: 199–212.
Venables W.N., Ripley B.D. 2002. Modern applied statistics with S. 4th ed. NewYork: *Springer*
Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. *Generalization in pollination systems, and why it matters. Ecology. **77.** 1043-1060.*
Weber A, Clark JL, Möller M. 2013. A new formal classification of Gesneriaceae. *Selbyana* **31**: 68–74.
Wessinger CA, Rausher MD. 2014. Ecological transition predictably associated with gene degeneration. *Molecular Biology and Evolution* **32**: 347–354.
Wilson P, Wolfe AD, Armbruster WS, Thomson JD. 2007. Constrained lability in floral evolution: counting convergent origins of hummingbird pollination in *Penstemon* and *Keckiella*. *New Phytologist* **176**: 883–890.
Woo VL, Funke MM, Smith JF, Lockhart PJ, Garnock-Jones PJ. **2011**. New World origins of Southwest Pacific Gesneriaceae : multiple movements across and within the south pacific. *International Journal of Plant Sciences* **172**: 434–457.

Chapter 3

Gesneriaceae: building the transcriptomic resources

Abstract

Despite the extensive phenotypic diversity that characterizes the Gesneriaceae family there is a lack of genomic resources to investigate the molecular basis of such diversity. We aimed to construct these resources using a comparative transcriptomic approach applied on six Gesneriaceae species. Illumina sequencing and *de novo* assembly of floral and vegetative samples were used to characterize the expression profiles, and generate multi-gene sequence data for the species sampled.

We obtained a total of 802 Gb of clean data, and produced the assembly of six *de novo* transcriptomes with an average of 200,000 transcripts per species. All *de novo* transcriptomes showed good quality metrics, with the presence of all eukaryotic core genes, and equally represented COG classifications between species. The orthologous search pipeline produced 8,847 one-to-one groups, with 48% of them annotated using BlastP and BlastX. Multiple sequence alignments were generated from the orthologous groups, and those can be used for further comparative analyses. Raw data from every RNA-Seq was mapped to the reference transcriptome of each species, producing the raw counts per library as the basis for the expression analyses. We incorporated all raw and processed data into a MySQL database that allows for multiple queries at the nucleotide, protein, annotation, and expression level.

This chapter provides the first step towards a comprehensive multi-species transcriptome characterization in the Gesneriaceae family. These resources are the basis for comparative analyses among the studied species, but also allow the investigation of multiple metabolic pathways with the addition of other plant groups. The next-generation resources that we generated for the Gesneriaceae family will provide valuable data for taxonomic, evolutionary and developmental studies in these non-model species.

Introduction

The evolutionary process of plant species formation is one of the central questions of this thesis work. The understanding of genetic, phenotypic and ecological divergences and their contribution to generate reproductive isolation is required for an advance in this topic. The literature from the last decade pinpointed the call for large genomic sequencing projects to better integrate the patterns of molecular, chromosomal, and epigenetic evolution, into plant speciation (Rieseberg & Wendel, 2004; Bomblies & Weigel, 2007). More recent reviews presented already the advances and future research on the identification of genes underlying reproductive barriers in plants (Lexer & Widmer, 2008; Rieseberg & Blackman, 2010). However, most of the organisms listed correspond to model plant species (such as species from the genera *Antirrhinum*, *Arabidopsis*, *Helianthus*, *Mimulus*, *Nicotiana*, *Petunia*, and *Solanum*) where candidate genes, functional analyses and complete lines of research have been developed. The investigation of plant speciation and other evolutionary questions in non-model plant species is motivated by the patterns of ecological and phenotypic differentiation present on those (Elmer & Meyer, 2011), but it represents a challenge due to the lack of genomic resources. This challenge starts however to disappear thanks to the advent of next-generation sequence technologies and the use of comparative frameworks (Blavet *et al.*, 2011; Lulin *et al.*, 2012).

One of the technologies that have revolutionized the study of gene expression is the whole-transcriptome sequencing, or RNA-seq (Wang *et al.*, 2009). This technique directly accesses most of the expressed protein-coding genes in a sample, and allows the investigation of differences in gene expression between conditions or populations. One of the major advantages of the RNA-seq is its ability to provide information on the gene expression, but also sequence data, without any previous knowledge of the biological system (i.e. a repertoire of genes as in microarrays or any genomic resources). The increased adoption of this technique to address ecological and evolutionary questions relies partially in its applicability to non-model organisms and its growing tool kit and bioinformatic support (Orsini *et al.*, 2013; Wolf, 2013). A critical step in the implementation of a RNA-seq study is the experimental design, sample size (replicates), and sequencing depth, which can all affect the power of detecting differential gene expression. However, the current developments have helped to propose rules and workflows that overcome these issues (Todd *et al.*, 2016).

The investigation of non-model species in a genomic context have shown a potential contribution to ecological and evolutionary studies (Ekblom & Galindo, 2011). Our previous findings highlight the value of the Gesneriaceae family for understanding plant speciation process in the Neotropics (see chapter 1 and 2). However, genomic or transcriptomic resources in the family are required. This family represents an ideal model to investigate plant-pollinator interactions, and the associated floral changes, due to the varied pollinator interactions and the convergent evolution of floral morphologies (Perret *et al.*, 2007; Clark *et al.*, 2012). The study of floral morphology is usually tackled by traditional genetic approaches to identify controlling genes (see review in Glover, 2014, section III). However, floral traits may evolve as a combination of long term evolutionary forces, such as recurrent selection from pollinators, herbivory, and genetic constraints and pleiotropy. Analyzing such complex trait evolution in a genomic context with intra- and inter-specific information has a great potential for identifying the floral genetic programs for plant-pollinator interactions (Clare *et al.*, 2013). The current genomic resources and the studies on the genetic control of flower morphologies in the family are scarce (Chiara *et al.*, 2013; Alexandre *et al.*, 2015). Here, we build novel genomic resources for six species within the family, following a “model-clade” approach with the characterization of multiple related species, making gene expression, sequence evolution, and identification of candidate genes, more accessible and robust to a large evolutionary scale (Chanderbali *et al.*, 2016). This chapter describes the methodological details required for the establishment of the Gesneriaceae transcriptomic resources, and preparation of data for further analyses.

The extraction of information, such as nucleotide and protein sequences for the assembled transcripts, annotation tables, and alignments, paves the way for understanding the genetic basis of distinct floral morphologies, evolved during pollination shifts, addressed in chapter 4, and future directions. Here, we report the species and sample selection, the laboratory procedures, sequencing and assembly pipelines, and the construction of a database to organize and make the genomic data accessible to a larger public.

Methods

Plant material

The Gesneriaceae family exhibits an outstanding flower diversity that can be related to repeated adaptations to different functional groups of pollinators such bees, hummingbirds and bats (Perret *et al.*, 2001; Martén-Rodríguez *et al.*, 2009). The high rate of transition between these pollination syndromes makes Gesneriaceae an ideal plant group to investigate the role of pollinator shifts in the evolution of flowers (see in chapter 2, and Fig. 1). Here, we selected six species based on their pollination types, representing three pairs of closely related species that have undergone pollination transitions (Fig. 1B). Three of them are bee-pollinated (*Sinningia eumorpha* = SE, *Paliavana tenuiflora* = PT, and *Nematanthus albus* = NA), whereas three others are hummingbird-pollinated (*Sinningia magnifica* = SM, *Vanhouttea calcarata* = VC, and *Nematanthus fritschii* = NF, see Table 1 for details and figures 1B and 3). All species are diploid, and distributed in the Brazilian Atlantic Forest (Fig. 2A), and were grown in a greenhouse at the Botanical Garden of Geneva in Switzerland (Conservatoire et Jardin botaniques de la Ville de Genève, CJB, Fig. 3).

Table 1. Information for the selected species and samples, accession numbers correspond to the material at the Conservatoire et Jardin Botanique de Gèneve (CJB).

Species name	Accession CJB	Pollinator	References
Nematanthus albus	AC20036937J AC1369	Bee	Wolowski, M pers.communication
Nematanthus fritschii	AC1105	Hummingbird (*Ramphodon naevius*)	Franco and Buzato (1992)
Sinningia eumorpha	200807305 20090391J0	Bee (*Bombus morio*, and other bees)	SanMartin-Gajardo and Sazima (2004)
Sinningia magnifica	AC3615 AC23105	Hummingbird (*Colibri serrirostris*)	de Vasconcelos and Lombardi (2001)
Paliavana tenuiflora	AC2352 20036853N0	Bee (*Bombus brevivillus*)	Ferreira and Viana (2010)
Vanhouttea calcarata	AC2203 AC2404	Hummingbird (*Leucochloris albicollis*)	SanMartin-Gajardo and Sazima (2005)

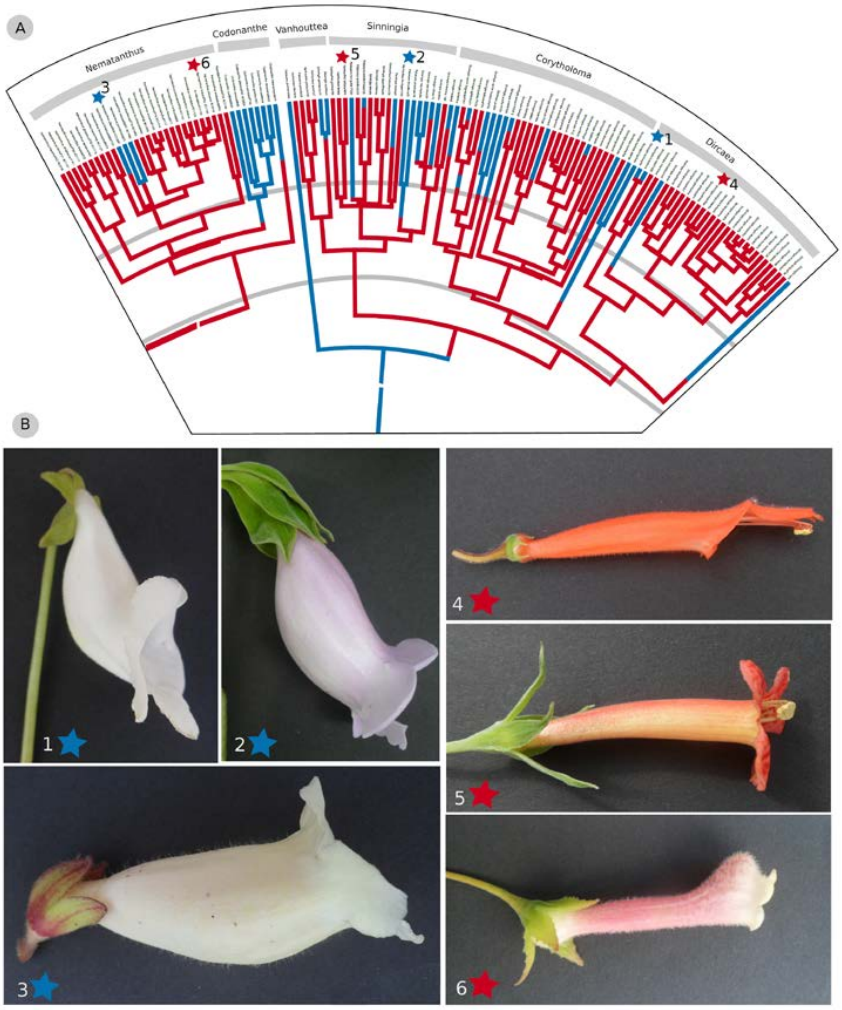


Figure 1. A) Evolution of pollination syndrome in the Neotropical Gesnerioideae showing the three selected pairs of species. This recontruction is a subset of the figure 1 in chapter 2. Blue and red branches correspond to insect and hummingbird pollination syndromes. Stars indicate the selected species. B) Species floral morphologies: 1, *S. eumorpha;* 2, *P. tenuiflora*; 3, *N. albus*; 4, *S. magnifica;* 5, *V. calcarata;* and 6, *N. fritschii.*

RNA extraction and libraries Illumina sequencing

Transcription patterns have a strong stochasticity component, and maximizing the understanding of biological variability within the budget of a project is a priority for RNA-seq experimental designs (Todd *et al.*, 2016). For this reason, we defined three developmental time points based on the percentage of the total flower size measured from the receptacle to the end of the corolla tube: bud1= 0-33%, bud2 = 33-75%, and flower 75-100%. We collected each species at every floral stage and vegetative material from two biological replicates represented by different accessions or individuals cultivated at the CJB (Table 1, following the design in Fig. 2B). Each sample contained at least two separate flowers to average the expression within a stage (see Appendix 1), and multiple floral tissues, such as sepals, the limb portion of the flower, flower tube, anthers including the filament, and the stigma including the style were all combined. All samples were collected from May 2013 to November 2014, and were immediately frozen in liquid nitrogen and stored at -80°C. Plants are grown in greenhouse homogeneous conditions, minimizing climatic variations between collecting days.

RNA was extracted with the Qiagen Rneasy Plant kit (cat. Nos. 74904) and treated with DnaseI (Qiagen) according to manufacturer's instructions. Illumina TruSeq stranded paired-end mRNA libraries were performed using 2µg of total RNA, following the library prep kit instructions (protocol version 15031047, Revision D, September 2012) for 300 bp fragments. Libraries were constructed by pairs of species, thus minimizing the batch effects when comparing phenotypes. Library concentration, integrity and size were determined with Agilent Fragment Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) and Qubit fluorometric quantitation (Thermo Fisher Scientific Inc). Illumina sequencing was performed with 100 cycles paired-end reads in a HiSeq 2500 at the Lausanne Genomic Technologies Facility.

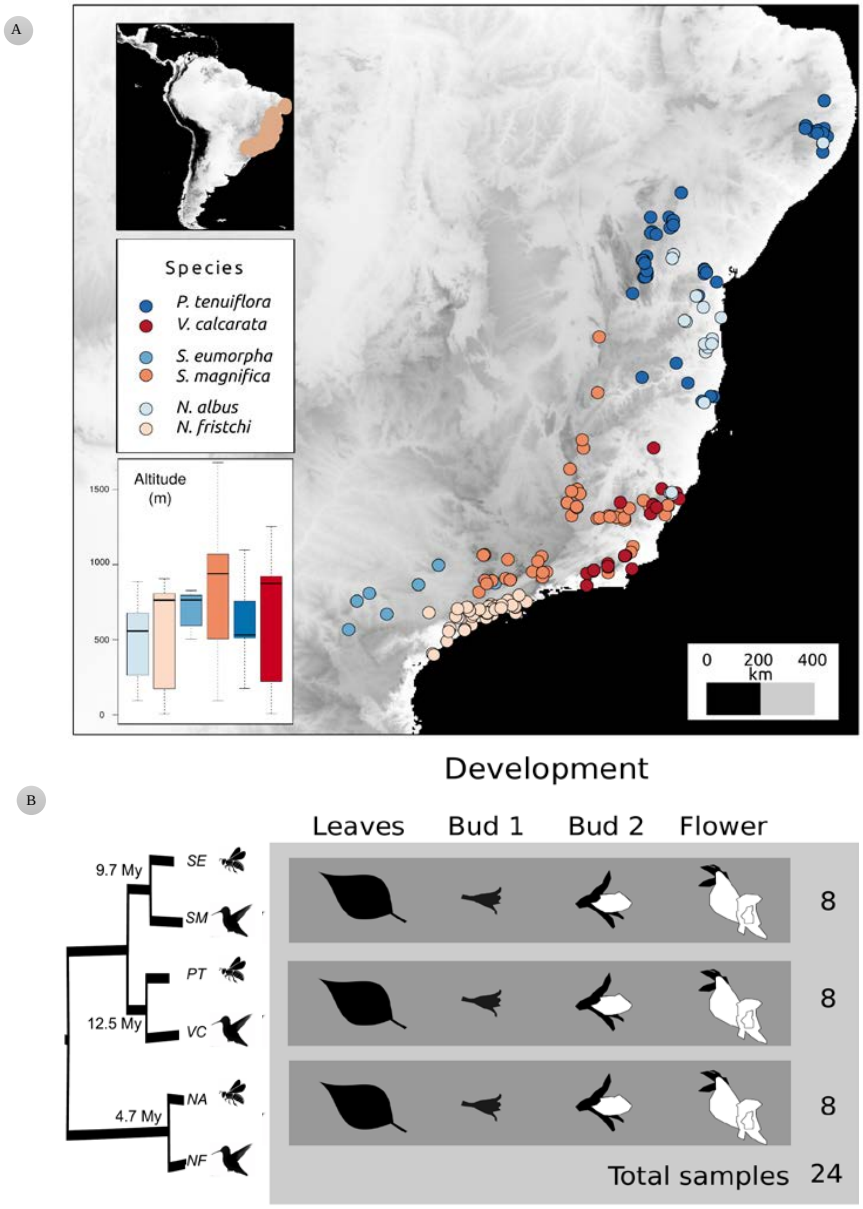


Figure 2. A) Geographical and altitudinal distribution of the six species in Brazil (Data from Perret *et al.*, 2006; Serrano-Serrano *et al.*, 2015). B) Experimental design for developmental and phylogenetic sampling. Mean divergence times between species pairs are presented on the tree nodes (Serrano-Serrano *et al.* submitted, see chapter 2).

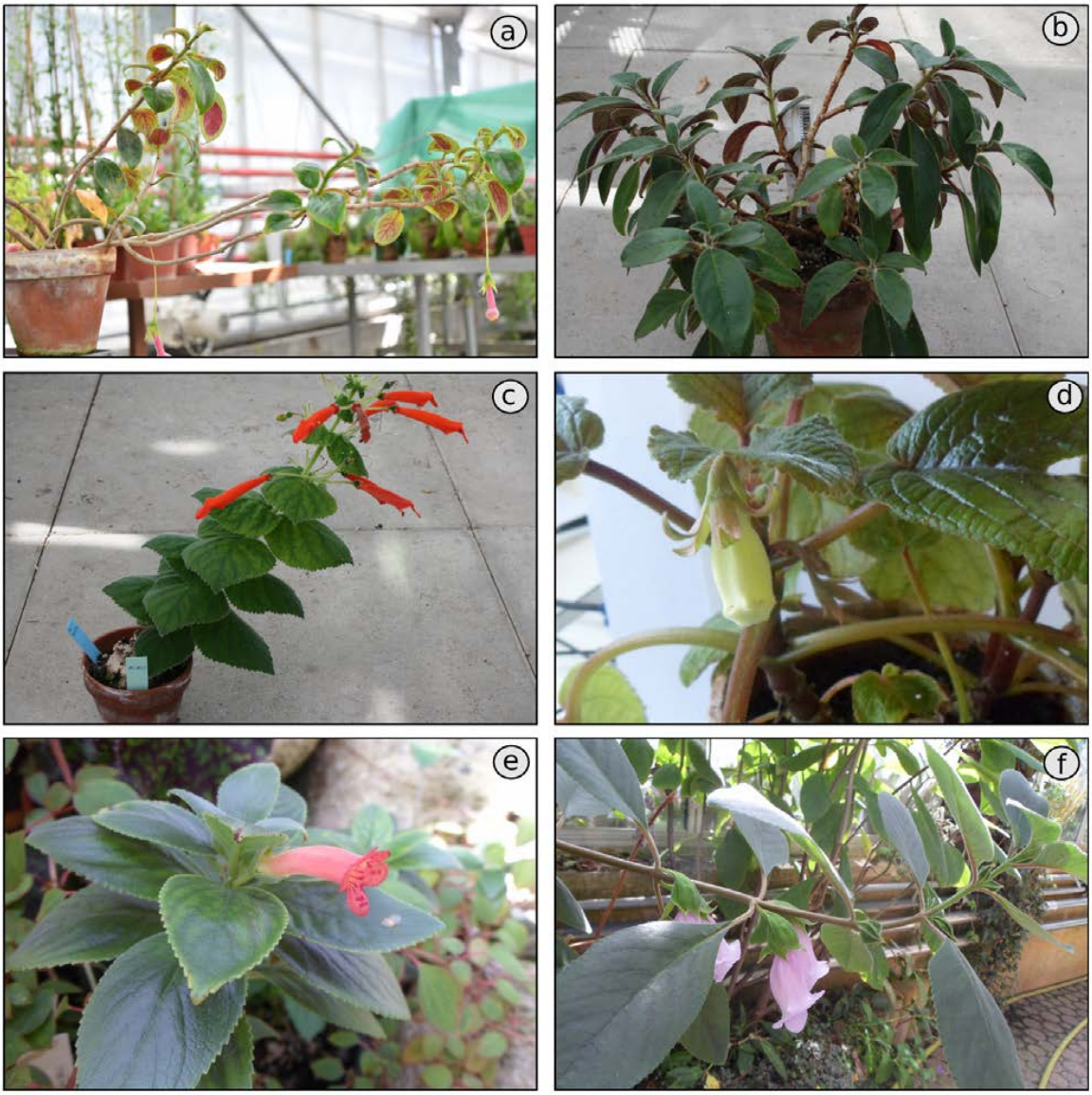


Figure 3. Vegetative morphologies for the selected species. a) *N. fritschii*, an epiphytic plant, with solitary and pendent resupinate flowers, with a gibbous corolla. Protandrous flowers, develop anthers and then stigma, over a period of 5 days. Leaves often with abaxial purple spots (Franco & Buzato, 1992). b) *N. albus*, an epiphytic fragrant plant, densely pilose leaves with trichomes, often reddish abaxial surfaces. Non-resupinate and axial flowers strongly fragrant, with a mixture of aldehydes (sweet and citric fragrance) identified (Chautems *et al.*, 2005). c) *S. magnifica*, plant with tuber habit, erect inflorescences, flowers with fusion and expansion of two dorsal corolla lobes (Chautems *et al.*, 2010). d) *S. eumorpha*, tuberous herb with a rosette habit and campanulate flowers (SanMartin-Gajardo & Sazima, 2004). e) *V. calcarata*, herbaceous plant occurring in rocky soils (SanMartin-Gajardo & Sazima, 2005). f) *P. tenuiflora*, shrub species occurring on rocky outcrops, flowers produce a large amount of nectar (Ferreira & Viana, 2010).

Transcriptome assembly, annotation

Raw reads were preprocessed, trimmed and filtered with a minimum length of 80 nucleotides and a quality score higher than 20, using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). *De*

novo transcriptomes for each species were assembled using the Trinity pipeline (Grabherr *et al.*, 2011, version 2.0.3) using a minimum contig length of 200. The large amount of raw reads from all libraries per species (flower stages and vegetative) was *in silico* normalized to a maximum coverage of 50. We filtered the lowest 5% of the transcript length distribution obtaining the final transcript set per species. Open reading frames (ORFs) were predicted with TransDecoder (Haas *et al.*, 2013).

All transcripts and ORFs were annotated using BlastX and BlastP against the SwissProt database. Sequence contaminants were screened and removed using the Blast information that matched with bacterial, fungal or any other non-plant genetic material. Blast annotations were filtered to avoid spurious hits using a threshold for the e-value and identity higher than $1x10^{-6}$ and 55%, respectively. Trinotate was used to integrate the functional annotation (http://trinotate.github.io), selecting one unique top blast hit and gene ontology (GO) annotation. Gene ontologies were plotted and compared between species using the WEGO webtool (http://wego.genomics.org.cn/cgi-bin/wego/index.pl).

Transcriptome quality checking

The assembled transcriptomes for every species were examined for quality and completeness using two measures: the Orthologous Hit Ratio (OHR, O'Neil *et al.*, 2010) and the Core Eukaryotic Genes Mapping Approach (CEGMA) analysis (http://korflab.ucdavis.edu/datasets/cegma/). The OHR is computed as the percentage of a gene in the transcriptome that matches a putative ortholog in tomato. It is calculated by dividing the length of the putative coding region by the total length of the orthologous gene (see Fig. 4). We performed a BlastX of our transcripts against the set of transcripts from tomato (*S. lycopersicum*, predicted protein database ITAG 2.4, 34'725 sequences on December 2014), and consider the best hit with an E-value < $1x10^{-6}$ to be **β** -orthologs. For the CEGMA analysis we used the COGs (Clusters of Orthologous Groups for Eukaryotes) to search the 458 highly conserved core proteins that matched our predicted ORFs (Parra *et al.*, 2007).

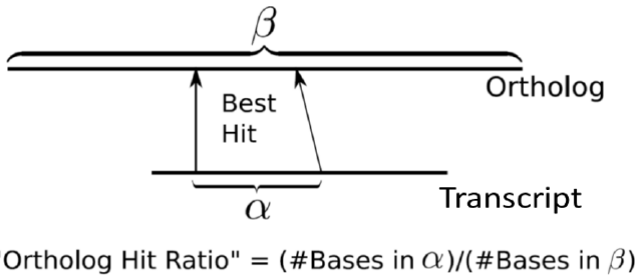


Figure 4. Description of OHR calculation. **α** correspond to the assembled transcripts, and the **β** to the orthologous genes in the reference (taken from O'Neil *et al.* 2010).

Orthologous search and alignments

OrthoMCL (version 2.0.9, Li *et al.*, 2003) was used for the identification of orthologous groups (OG) between the six species. This step is required for all comparative analyses, and uses the predicted proteins from every transcriptome (ORFs) to conducted all-against-all Blast searches. We used the information of all constructed OG to select the 1:1 orthologous groups, which correspond to genes represented by a single copy per species. Every OG was annotated using the Blast information from each associated transcript.

Predicted protein sequences from all species belonging to every OG were aligned using MAFFT (version 7.187, Katoh & Standley, 2013). Protein alignments were converted into codon alignments using command line Pal2Nal script (Suyama *et al.*, 2006). Cases of multiple transcript sequences from a single species were handled using the EMBOSS consensus calculator (EMBOSSCons, http://emboss.sourceforge.net/apps/cvs/ emboss/apps/cons.html). This tool produced a single sequence representing the consensus sequence for the gene of the species.

Mapping reads for expression analysis

Cleaned RNA-Seq reads from each library were mapped into all species transcript sets using the RSEM pipeline (Li & Dewey, 2011). This gene expression quantification involves two steps, the first one aligns the reads to the reference transcripts, using Bowtie (Langmead *et al.*, 2009), while supporting paired-end and strand-specific reads. The second step, estimates the gene/isoform abundances via maximum likelihood estimates to produce expected-counts (Expectation-Maximization, ME algorithm). They are called expected as it uses all aligned reads, even those that do not map uniquely to a single transcript. This procedure gives the expression values for the full gene length that are further normalized between species. This step was performed with *align_and_estimate_abundance.pl* script available from the Trinity tools.

Database construction

We used a MySQL implementation to generate a structured database with the different types of information generated. The data was incorporated by using a customized Python script developed by Marion Patxot Bertran, as part of her first step Master project (MLS MSc program at UNIL during September to December, 2015).

Table 2. RNA-seq library concentration for the 48 samples sequenced.

	Library concentration [ng/ul]			
Species	Bud 1	Bud 2	Flower	Vegetative
N. albus	51.0	19.3	27.4	19.1
	51.6	50.0	50.0	17.1
N. fritschii	19.4	14.1	9.94	16.7
	57.4	40.3	49.8	8.28
S. eumorpha	13.5	28.2	18.5	9.31
	6.0	6.61	6.9	8.06
S. magnifica	28.8	14.9	21.0	18.6
	6.83	7.19	3.85	29.0
P. tenuiflora	18.7	20.0	19.2	20.2
	52.5	52.5	56.3	11.6
V. calcarata	28.2	22.8	21.0	8.01
	54.9	59.6	53.0	16.1

Results and discussion

Plant material, RNA extraction and library construction

We sampled all species under the developmental schema proposed (see figure 2B, and appendix 1), and successfully extracted RNA with variable sample concentration. Extraction quantities and

qualities were suitable for the preparation of Illumina libraries (Table 2). We generated Illumina raw reads with a similar sequencing effort between species and samples (280 to 332 million reads, Table 3).

Table 3. Illumina generated reads for each library and species. R1 and R2 correspond to the biological replicates.

Stage	NA	NF	SE	SM	PT	VC
Bud 1 R1	39,167,623	36,702,305	38,107,356	34,642,950	37,459,616	34,144,218
Bud 1 R2	28,237,393	43,179,065	58,915,805	36,269,381	28,961,714	37,567,360
Bud 2 R1	38,366,401	33,898,191	33,082,347	32,940,434	37,773,085	37,010,683
Bud 2 R2	41,822,365	60,268,340	47,344,551	66,952,370	41,021,858	37,001,832
Flower R1	42,810,862	39,580,742	43,492,497	36,723,815	35,331,819	42,490,378
Flower R2	34,053,850	77,227,386	42,061,213	45,995,005	29,704,332	32,797,914
Vegetative R1	26,780,468	30,565,766	22,566,904	38,932,470	48,905,586	28,236,772
Vegetative R2	37,148,495	26,541,793	30,905,818	40,473,831	33,644,447	31,238,710
Total reads	**288,387,457**	**347,963,588**	**316,476,491**	**332,930,256**	**292,802,457**	**280,487,867**

Transcriptome assembly, annotation

The assembly statistics are presented in Table 4. The number of genes between species was very similar, except for NF and SE species that have a slightly higher number of genes (and transcripts). The annotation of the six transcriptomes showed that only a low proportion of transcripts have significant Blast matches with existing protein and nucleotide sequences (12.38% and 19.07% for BlastP and BlastX, respectively). This pattern is common for non-model species and *de novo* assemblies where blast tools are known to fail the annotation of around 75 % of genes to any other known organism (Chiara *et al.*, 2013; DeBiasse & Kelly, 2016). Gene ontologies indicated that the proportion of assembled genes associated to each functional category is similar for all species datasets, and all *de novo* assemblies are likely comparable (Fig. 5). High and low abundance terms are shared between species, indicating that transcriptome assemblies have very similar compositions.

Transcriptome quality checking

The evaluation of transcriptome completeness performed using the CEGMA method showed a percentage between 99.6 and 100.0% of the Core Eukaryotic Genes mapped to the generated transcripts. Additionally, the evaluation of OHR indicated that the assembled genes covered a large proportion of the putative reference orthologs in tomato (Fig. 6). The estimated proportion of genes with a high overlap with the reference (larger to 0.8) ranged between 50.9 to 59.3% for all the species.

Table 4. Summary of the assembly statistics for the six Gesneriaceae species using Trinity.

Species	No. genes	No. transcripts	GC %	Median contig length	Average contig length	No. ORFs
NA	83,914	193,438	38.52	1,285	1,748	156,528
NF	104,760	233,230	38.47	1,000	1,548	171,808
SE	112,640	218,032	38.51	1,179	1,687	171,889
SM	75,322	179,888	38.31	1,089	1,545	131,000
PT	95,761	193,005	38.47	1,054	1,565	147,010
VC	85,031	185,092	38.53	1,032	1,525	136,772

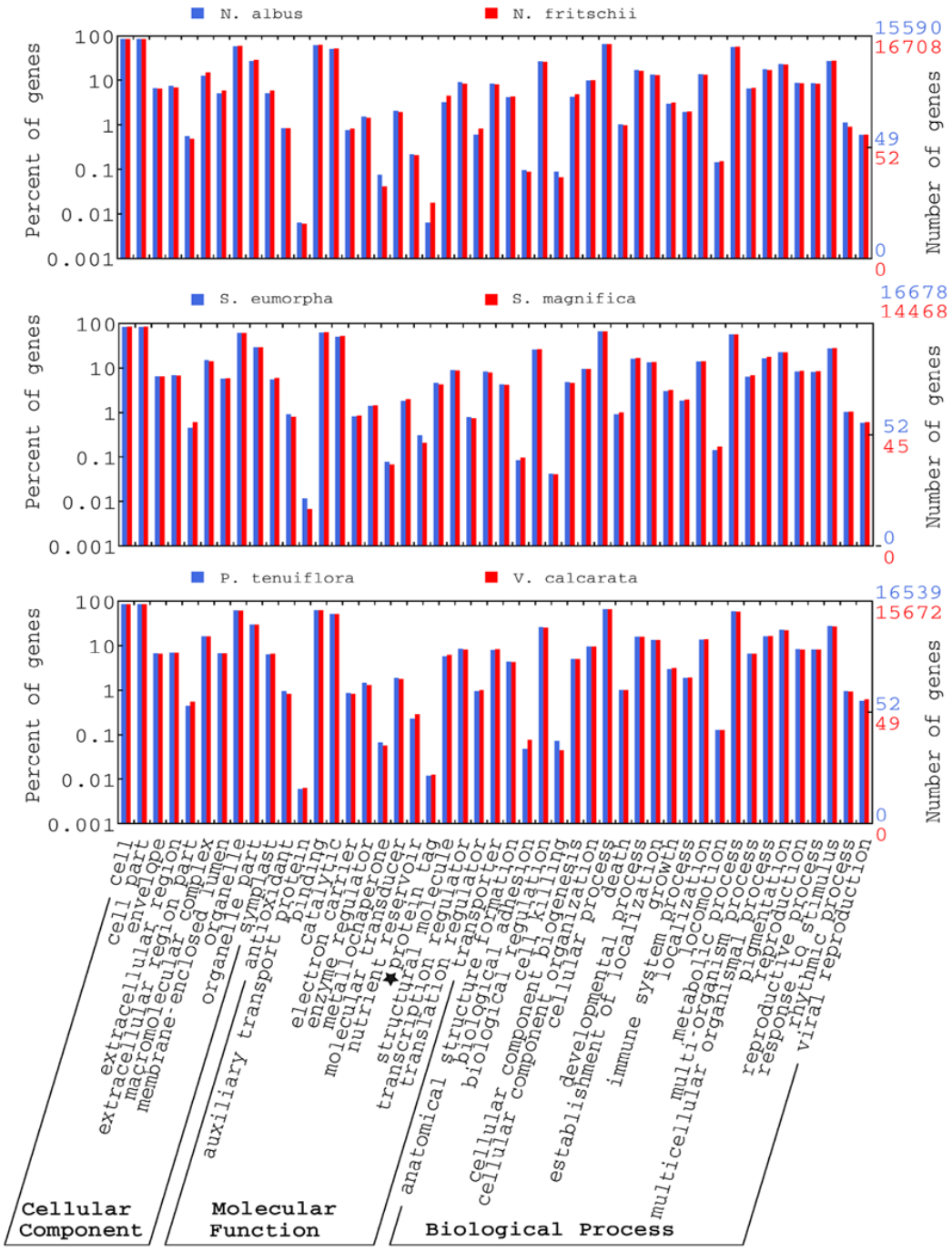


Figure 5. Histogram of GO classifications for each pair of species. The WEGO plot show the three main ontology categories: Cellular component, molecular function, and biological process. Black star shows a functional category that is missing in the pair SE - SM.

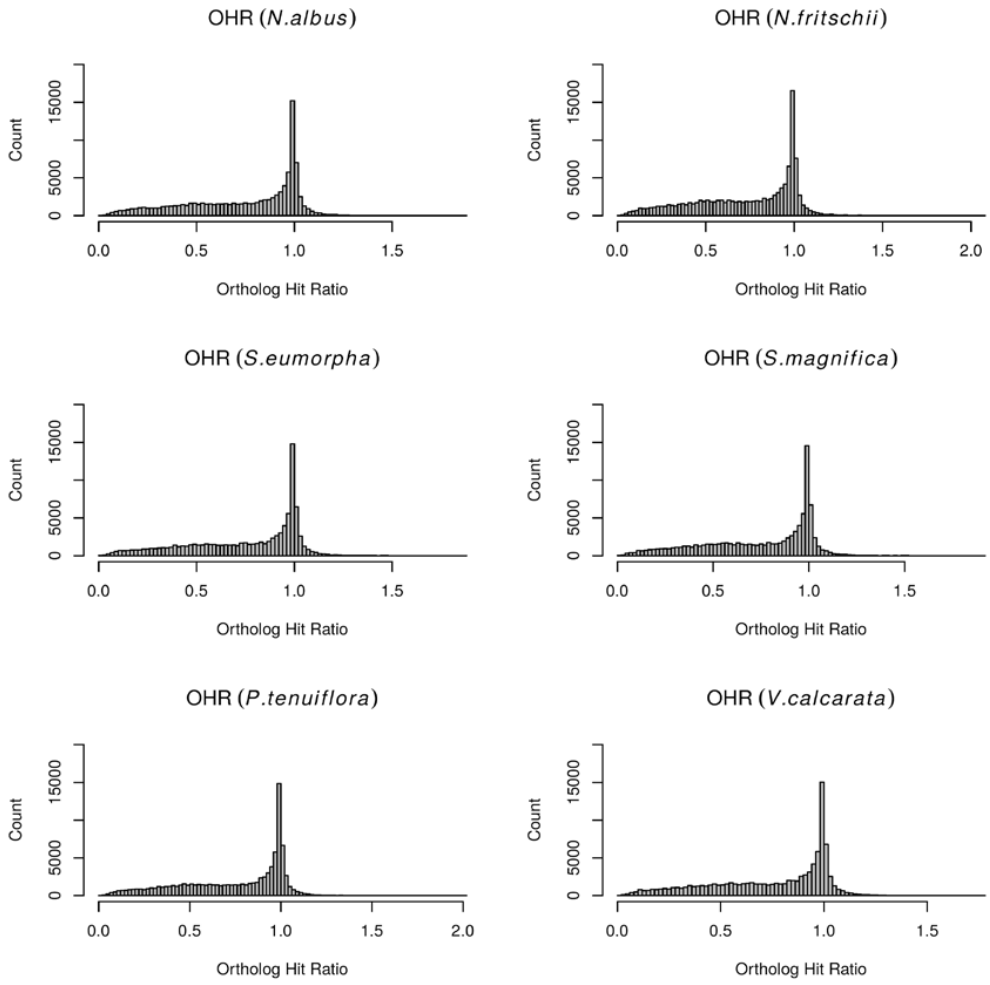


Figure 6. Distribution of orthologous hit ratio for all species genes. A ratio of 1.0 indicates that the gene has likely the same length than the reference orthologs (from tomato). The percentage of genes with an overlap larger than 0.8 is NA= 53.9%, NF = 59.3, SE = 51.4, SM = 51.3, PT = 50.9, VC = 53.7, respectively.

Orthologous search and alignments

OrthoMCL results provided the potential groups of orthologous genes between the six species. The number of sequences in each group varied from two to a few hundred, with most of the orthologous groups composed by less than 14 sequences (Fig. 7). For the downstream analyses we identified 8848 one-to-one orthologous groups between the six species, with around a 48% of those (4220 OGs) functionally annotated. This amount of OGs is comparable with other surveys between closely related species (Zhang *et al.*, 2013). We produced a multiple sequence alignment (MSA) for every OG as the basis for downstream analyses (see Fig. 8).

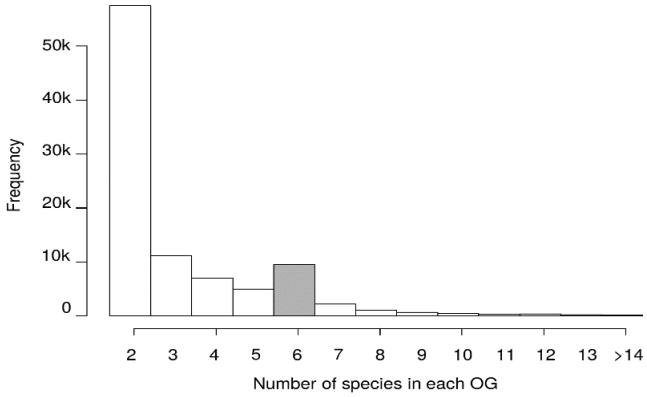


Figure 7. Frequency distribution for OGs with the different number of sequences per group. Groups with more than 14 species were removed from the plot (total of 879 OGs). Gray bar represents the one-to-one orthologous groups.

Mapping reads for expression analysis

RSEM mapping produced expected expression values for all the genes in every library for developmental time and species conditions. All conditions were integrated to generate the multi-species matrix abundance. This matrix was filtered to contain only the expression for the one-to-one OG, and constitutes the raw expression data for all further analyses, such as differential gene expression (see Fig. 8).

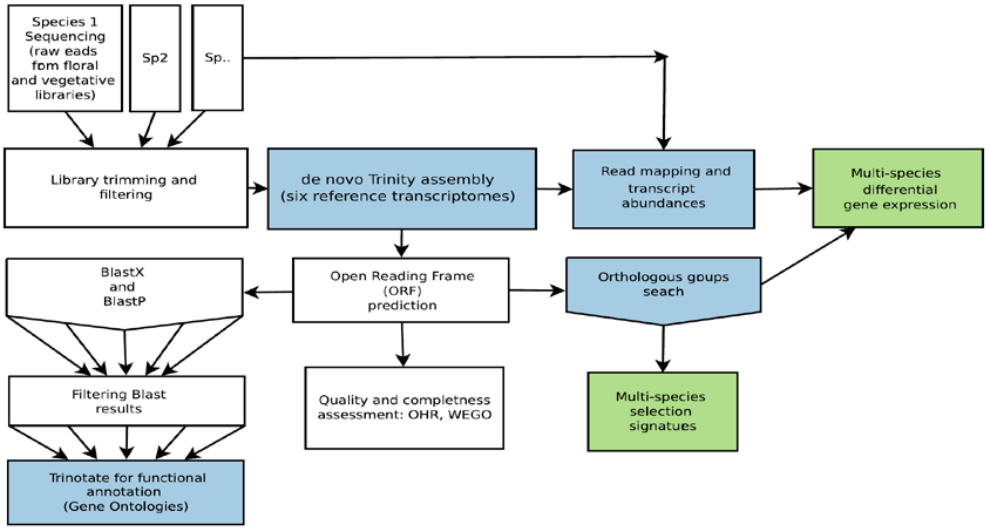


Figure 8. Bioinformatic pipeline for the construction of transcriptomic resources in the Gesnerioideae species. White boxes correspond to data processing analyses performed, blue boxes to data sets generated, and green boxes to downstream analyses (see Chapter 4). This pipeline was designed and implemented by ML Serrano-Serrano and Anna Marcionetti, as part of her Master project (MLS MSc program at UNIL during 2014-2015).

Database construction

The Gesneriaceae database relied on four types of generated data (see Fig. 8), the raw transcriptome assembly with the ORF prediction, the functional annotation data, the transcript/gene abundances, and the orthologous groups. Relational connections between the generated tables allow the search and extraction of information (Fig. 9). The database is stored at sftp://srvphylo.unil.ch, and can be directly queried and explored, but may require further developments for public access.

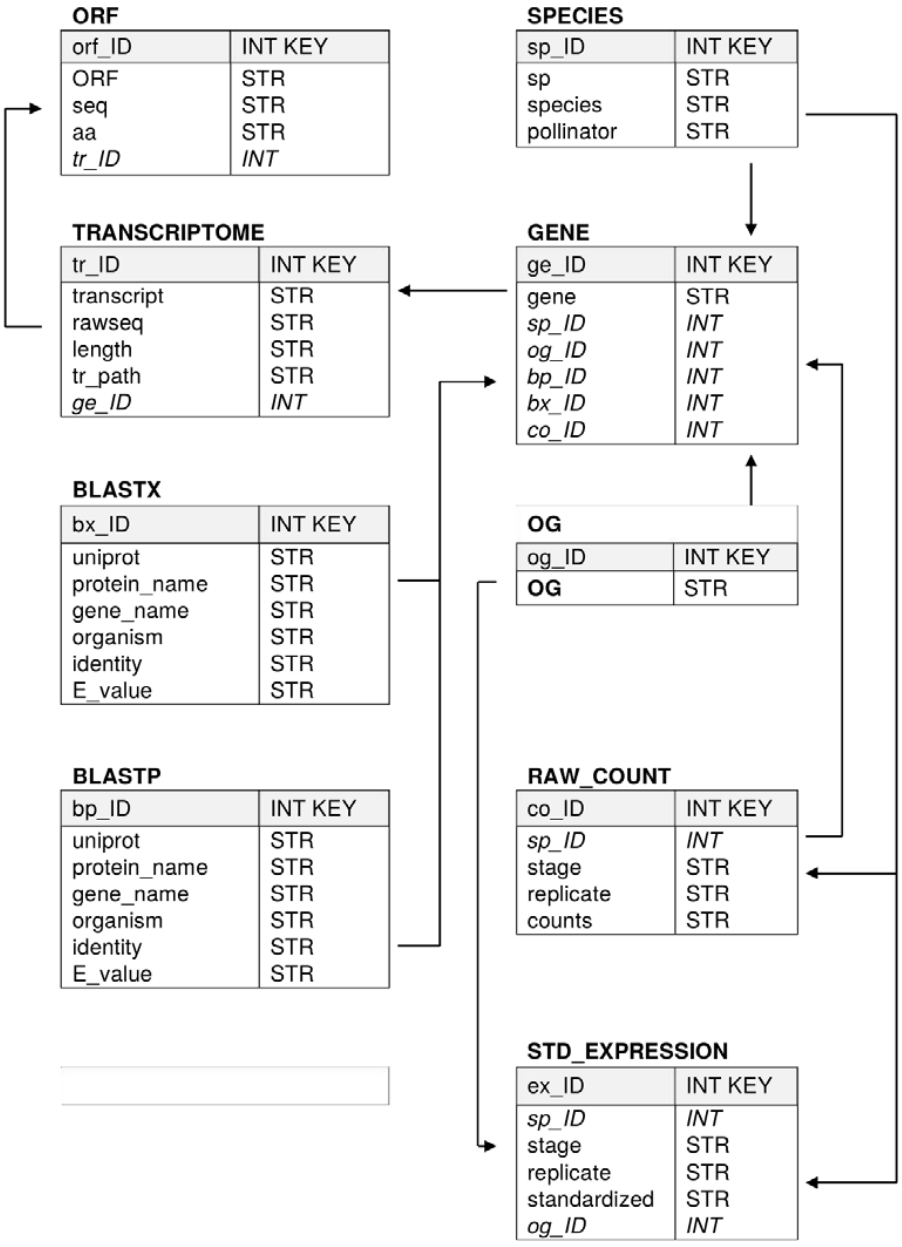


Figure 9. MySQL database structure for Gesnerioideae transcriptomic resources.

Conclusions

The six transcriptomes described here generated about 150.000 genes per species and 8000 OGs between them. The quality and quantity of data is comparable with other studies, and will improve the available genomic/transcriptomic data for the whole Gesneriaceae family. Our results provided large-scale sequence data for the six related species (only 6 nuclear genes were available for the family previous to this work), and expression level data for each of those, in three floral developmental stages and vegetative material. These resources will facilitate the investigation of ecological and evolutionary questions within the *Sinningia* and *Nematanthus* genera. The transcriptomes assembled will allow to test whether the similar floral morphologies between the studied species will share patterns of gene expression

and molecular evolution. Additionally, the developed resources constitute the opening material for designing probes and primers for phylogenetic and population studies, as well as, the basis for the identification of candidate genes, and further complementation of any genomic or experimental survey. The results in this chapter make multiple Gesneriaceae species accessible for comparative analyses at a larger macro-evolutionary scale by incorporating additional information from model plant species, or any other species currently available.

References

Alexandre H, Vrignaud J, Mangin B, Joly S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus Rhytidophyllum (Gesneriaceae). *PeerJ* **3**: e1028.

Blavet N, Charif D, Oger-Desfeux C, Marais G, Widmer A. 2011. Comparative high-throughput transcriptome sequencing and development of SiESTa, the Silene EST annotation database. *BMC Genomics* **12**: 376.

Bomblies K, Weigel D. 2007. Arabidopsis: a model genus for speciation. *Curr Opin Genet Dev* **17**(6): 500-504.

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE. 2016. Evolving Ideas on the Origin and Evolution of Flowers: New Perspectives in the Genomic Era. *Genetics* **202**(4): 1255-1265.

Chautems A, Lopes TCC, Peixoto M, Rossini J. 2005. Five new species of Nematanthus Schrad.(Gesneriaceae) from eastern Brazil with a revised key to the genus. *Selbyana*: 210-224.

Chautems A, Lopes TCC, Peixoto M, Rossini J. 2010. Taxonomic Revision of Sinningia Nees (Gesneriaceae) IV: Six New Species from Brazil and a Long Overlooked Taxon. *Candollea* **65**(2): 241-266.

Chiara M, Horner DS, Spada A. 2013. De novo assembly of the transcriptome of the non-model plant Streptocarpus rexii employing a novel heuristic to recover locus-specific transcript clusters. *PLoS One* **8**(12): e80961.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome biology* **14**(6): 1.

Clark JL, Funke MM, Duffy AM, Smith JF. 2012. Phylogeny of a Neotropical Clade in the Gesneriaceae: More Tales of Convergent Evolution. *International Journal of Plant Sciences* **173**(8): 894-916.

de Vasconcelos MF, Lombardi JA. 2001. *Hummingbirds and their flowers in the campos rupestres of southern Espinhaço Range, Brazil*: Melopsittacus publicaçãoes científicas.

DeBiasse MB, Kelly MW. 2016. Plastic and Evolved Responses to Global Change: What Can We Learn from Comparative Transcriptomics? *J Hered* **107**(1): 71-81.

Ekblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* **107**(1): 1-15.

Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol Evol* **26**(6): 298-306.

Ferreira PA, Viana BF. 2010. Pollination biology of Paliavana tenuiflora (Gesneriaceae: Sinningeae) in Northeastern Brazil. *Acta Botanica Brasilica* **24**(4): 972-977.

Franco A, Buzato S. 1992. Biologia floral de Nematanthus fritschii (Gesneriaceae).(Floral biology of Nematanthus fritschii (Gesneriaceae).). *Rev. Brasil. Biol* **52**(4): 661-666.

Glover B. 2014. *Understanding Flowers and Flowering Second Edition*: OUP Oxford.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature biotechnology* **29**(7): 644.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* **8**(8): 1494-1512.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**(4): 772-780.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**(3): 1.

Lexer C, Widmer A. 2008. Review. The genic view of plant speciation: recent progress and emerging questions. *Philos Trans R Soc Lond B Biol Sci* **363**(1506): 3023-3036.
Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.
Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research* **13**(9): 2178-2189.
Lulin H, Xiao Y, Pei S, Wen T, Shangqin H. 2012. The first Illumina-based de novo transcriptome sequencing and analysis of safflower flowers. *PLoS One* **7**(6): e38653.
Martén-Rodríguez S, Almarales-Castro A, Fenster CB. 2009. Evaluation of pollination syndromes in Antillean Gesneriaceae: evidence for bat, hummingbird and generalized flowers. *Journal of Ecology* **97**(2): 348-359.
O'Neil S, Dzurisin JD, Carmichael RD, Lobo NF, Emrich SJ, Hellmann JJ. 2010. Population-level transcriptome sequencing of nonmodel organisms Erynnis propertius and Papilio zelicaon. *BMC Genomics* **11**(1): 1.
Orsini L, Andrew R, Eizaguirre C. 2013. Evolutionary Ecological Genomics. *Molecular Ecology* **22**: 527-531.
Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**(9): 1061-1067.
Perret M, Chautems A, Spichiger R. 2006. Dipersal-vicariance analyses in the tribe Sinningieae (Gesneriaceae): a clue to understanding biogeographical history of the Brazilian Atlantic forest *Annals of the Missouri Botanical Garden* **93**(2): 340-358.
Perret M, Chautems A, Spichiger R, Barraclough TG, Savolainen V. 2007. The geographical pattern of speciation and floral diversification in the neotropics: the tribe Sinningieae (Gesneriaceae) as a case study. *Evolution* **61**(7): 1641-1660.
Perret M, Chautems A, Spichiger R, Peixoto M, Savolainen V. 2001. Nectar sugar composition in relation to pollination syndromes in Sinningieae (Gesneriaceae). *Annals of Botany* **87**(2): 267-273.
Rieseberg LH, Blackman BK. 2010. Speciation genes in plants. *Ann Bot* **106**(3): 439-455.
Rieseberg LH, Wendel JF. 2004. Plant speciation - Rise of the poor cousins. *New Phytol* **161**: 1-21.
SanMartin-Gajardo I, Sazima M. 2004. Non-Euglossine bees also function as pollinators of *Sinningia* species (Gesneriaceae) in southeastern brazil. *Plant Biol (Stuttg)* **6**(4): 506-512.
SanMartin-Gajardo I, Sazima M. 2005. Espécies de *Vanhouttea* Lem. e *Sinningia* Nees (Gesneriaceae) polinizadas por beija-flores: interações relacionadas ao hábitat da planta e ao néctar. *Revista Brasileira de Botânica* **28**(3): 441-450.
Serrano-Serrano ML, Perret M, Guignard M, Chautems A, Silvestro D, Salamin N. 2015. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evol Biol* **15**: 247.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic acids research* **34**(suppl 2): W609-W612.
Todd EV, Black MA, Gemmell NJ. 2016. The power and promise of RNA-seq in ecology and evolution. *Mol Ecol* **25**(6): 1224-1241.
Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**: 57-63.
Wolf JB. 2013. Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mol Ecol Resour* **13**(4): 559-572.
Zhang L, Yan H-F, Wu W, Yu H, Ge X-J. 2013. Comparative transcriptome analysis and marker development of two closely related Primrose species (Primula poissonii and Primula wilsonii). *BMC Genomics* **14**(1): 1.

Appendix 1A. Sampling of *N. albus* (left) and *N. fritschii* (right), in the three developmental stages: B1= Bud 1, B2= Bud 2, and FL= adult flower. White horizontal bar is a 1 cm scale.

Appendix 1B. Sampling of *P. tenuiflora* (left) and *V. calcarata* (right), in the three developmental stages: B1= Bud 1, B2= Bud 2, and FL= adult flower. White horizontal bar is a 1 cm scale.

Appendix 1C. Sampling of *S. eumorpha* (left) and *S. magnifica* (right), in the three developmental stages: B1= Bud 1, B2= Bud 2, and FL= adult flower. White horizontal bar is a 1 cm scale.

Chapter 4

Transcriptomic evidence for the parallel evolution of pollination syndromes in Gesneriaceae

Abstract

Shifts in pollinator types have occurred repeatedly during the diversification of the Neotropical Gesneriaceae species, producing highly convergent floral morphologies. The genes underlying these morphological changes have been studied from model species, but clade-models and comparative approaches are needed for an investigation of the concerted floral changes involved in pollinator shifts. We compared the transcriptomes of six related species within the *Sinningia* and *Nematanthus* genera, which represent three independent pollinator shifts between bee and hummingbird functional groups. Here, we showed that gene expression profiles and sequence evolution differed between pollination types, with a very low proportion of concerted genetic changes. We specifically discuss changes in gene expression during flower development and signatures of positive selection, with a larger number of genes showing concerted selection signatures than differences in expression. We propose a series of candidate genes affecting floral shape, size and color, among other correlated traits with pollinator attraction that provide the basis for further functional experiments. Our results reflected that multiple genetic routes could produce similar floral morphologies suitable for hummingbird pollination. These alternative mechanisms may have facilitated the highly labile evolution of pollination systems in this clade of Neotropical plants.

Introduction

Evolutionary convergence describes the independent evolution of similar phenotypes different lineages. The molecular basis of these predictable convergent solutions is determined by genetic constraints, phylogenetic and population histories, and natural selection (Rosenblum *et al.*, 2014). Instances of repeated evolution have been of recent interest in evolutionary biology, especially to investigate whether phenotypic convergences are linked to molecular convergences (Steiner *et al.*, 2009; Foote *et al.*, 2015). Non-model organism are ecologically interesting and represent a potential source of convergent phenotypes, fortunately the advent of next-generation sequencing technologies make more feasible the investigation of their genomic basis (Elmer & Meyer, 2011). For instance, comparative transcriptomic surveys have helped to identify potential genes associated with complex traits such as convergent bioluminescent organs in squids (Pankey *et al.*, 2014), and eusociality in insects (Woodard *et al.*, 2011; Berens *et al.*, 2015). However, the extent of phenotypic convergence, the nature and reuse of the genetic mechanisms is still a challenge for many traits and lineages (Martin & Orgogozo, 2013; Ord & Summers, 2015).

Floral morphologies are an impressive example of phenotypic convergence at small and large taxonomic scales (Brown & Kodric-Brown, 1979; Schemske, 1981; Weiss, 1995). First, their genetic architecture seems to be similar, encompassing a series of shared homeotic and highly customizable genes (Pires & Dolan, 2012). Second, the selection exerted by pollinators on independent plant lineages seems to be concerted and adaptive (Thomson & Wilson, 2008; Harder & Johnson, 2009), though it is not an exclusive agent of selection or evolutionary force (Strauss & Whittall, 2006). This concerted set of traits associated with the attraction of particular groups of pollinators is known as pollination syndromes, and include phenotypes to attract and reward pollinators according to their specific needs and behaviors (Fenster *et al.*, 2004; Rosas-Guerrero *et al.*, 2014).

Two groups of plant-pollinator interactions have been largely studied: insect-pollinated and bird-pollinated plants. Because these two types are analyzed here, a brief description of each of group, and their characteristic set of traits is provided below. Insect-pollination is an ancient ecological association, which can be traced back through the fossil record (Michez *et al.*, 2012). Insect-pollination is an inclusive category showing divergent selective pressures from diverse functional groups of insects (long-tongued bees, short-tongued bees, other Hymenoptera, Diptera, Coleoptera, Lepidoptera, Hemiptera, Neuroptera, see Kevan & Baker, 1983). Here, we will focus on bee-pollination syndrome, which include traits such as corolla landing platform, concentrated nectar, nectar guides, blue or yellowish color and fragrance that aim to attract bees (Fenster *et al.*, 2004; Glover, 2014). In contrast, bird-pollination seems to be a derived condition in many plant groups, and existing traits in insect-pollination, such as zygomorphy and tubular corollas, may have facilitated the transitions to bird-pollinated flowers (Cronk & Ojeda, 2008).

Hummingbird-pollination is a type of bird-pollination exclusive to the New World. Studies in *Penstemon* species (Plantaginaceae) indicated that floral morphologies are characterized by pro-bird and anti-bee traits (Castellanos *et al.*, 2004). These traits include organ exsertion (reducing pollen deposition and stigma contact for bees), lipless and constricted corollas (lacking a landing platform and making nectar less accessible for bees, or increasing handling time for hummingbirds), pendent flowers (reducing bee visitation). Additional major phenotypic convergences of the evolution of hummingbird pollination are reddish corollas, diluted nectar and lack of scent (Cronk and Ojeda 2008).

Evolutionary transitions between pollination types are recurrent in angiosperms, among closely related species or involving complete genera. Most of the groups investigated for pollination shifts have shown multiple acquisitions of hummingbird pollination, but few or none reversions to bees (Kay & Schemske,

2003; Wilson *et al.*, 2007; Alcantara & Lohmann, 2010). However, this dead-end pattern is escaped in some lineages (Whittall & Hodges, 2007; Tripp & Manos, 2008, see chapter 2) and attracts the curiosity about the level of phenotypic convergence, and the reversibility of the genetic changes required for pollination shifts. The examination of the genetic architecture associated to pollination syndrome traits have progressed in the recent years, especially by the usage of model systems such as *Antirrhinum, Ipomoea, Penstemon, Petunia* and *Mimulus* species (Bradshaw & Schemske, 2003; Zufall & Rausher, 2003; Stuurman *et al.*, 2004; Perez-Rodriguez *et al.*, 2005; Wessinger *et al.*, 2014). The current knowledge indicates that certain traits are controlled by a rather small number of genes or Quantitative Trait Loci (QTLs, such for color and scent), while others are more complex and may involve a large number of loci (i.e. floral size and shape, stamen and pistil sizes, nectar volume and composition). For two detailed reviews see Galliot *et al.* (2006) and Hermann and Kuhlemeier (2011).

Given the similarities in floral morphologies within a pollination syndrome, we explore the extent of concerted genetic changes occurring during pollination transitions. By concerted we mean the genes that respond in the same way to the three pollination transitions investigated. Recent evidence for the parallel evolution of hummingbird-pollinated flowers in *Ipomea* suggested that concerted changes have occurred at the developmental and genetic level (Des Marais & Rausher, 2010). Those changes are specifically affecting the production of pelargonidin-based pigments by down-regulation and *cis*-regulatory mutations in the genes involved in the anthocyanin pathway. However, the repeatability of these mechanisms have been scarcely tested, and mainly associated with floral color transitions (Smith *et al.*, 2013; Wessinger & Rausher, 2014), and many other traits remain to be investigated. These studies motivate the discovery of the genetic mechanisms, discriminating whether differences in gene expression or coding sequences are more relevant for changes in floral morphologies, and consequently generating reproductive isolation and speciation in plants (Pavey *et al.*, 2010; Butlin *et al.*, 2012).

Recent molecular evidence indicated the repeated origin of zygomorphic flowers, and shifts in pollinators, in the Malpighiaceae family is associated with divergences in gene expression and loss of function in the CYCLOIDEAE2-like transcription factors (Zhang *et al.*, 2012). The large taxonomic scale of this study is attractive, however it explored only the genetic responses associated to this transcription factor family, and additional genomic aspects, potentially relevant, are missing. Ideally, rather than focus in single gene families, the complete genome or transcriptome evaluation should provide a wider examination of the genetic signatures and adaptive changes during the evolution of a convergent phenotype. A fascinating example is provided by the study of the convergent evolution of eusocial insects, where transcriptomes brought evidence of substantial molecular convergence at gene pathways, rather than at the exact genes, and accelerated evolution in specific biological functions (Woodard *et al.*, 2011; Berens *et al.*, 2015). These patterns allow us to link specific genetic changes (gene expression, substitution rates, and selection signatures) with the origination of shared traits in independent lineages.

Here, we evaluate the overall concerted genetic mechanisms during pollination transitions, by comparing the transcriptomic information from multiple evolutionary replicates. Those replicates are sampled within the Gesnerioideaes subfamily, which comprises a large diversity of floral morphologies and convergences in pollination syndromes (Perret *et al.*, 2007; Marten-Rodriguez *et al.*, 2010). Recent macro-evolutionary analyses in the subfamily have shown that pollination syndrome transitions have evolved multiple times independently in the Neotropics (Serrano-Serrano *et al.* submitted). The high rates of transition between pollination syndromes, as well as the reversibility of hummingbird-pollination make the subfamily an ideal model to study the molecular genetic basis of pollination syndrome. This chapter examines the extent of concerted changes in gene expression, and molecular signatures of selection for three evolutionary shifts between bee and hummingbird pollination. We predict a large proportion of gene regulatory network elements underlying floral diversity, and aim to identify candidate genes for further

functional investigation. No previous study has attempted to characterize these coordinated changes associated with pollination syndromes, in a multi-species framework or "model clade" approach (though some examples for the evolution of zygomorphy, see Chanderbali *et al.*, 2016). This multi-species examination is valuable to understand the plant-pollinator floral evolution in a moderate timescale (Clare *et al.*, 2013). The great advantage of this approach rely on the potential generalization of the genetic basis of pollination syndromes at moderate evolutionary timescale. To achieve this, we built the first transcriptomic resources for six non-model related species (*N. albus, N. fritschii, S. eumorpha, S. magnifica, P. tenuiflora* and *V. calcarata*) with bee- and hummingbird pollination syndromes (laboratory and bioinformatics details are described in chapter 3).

Methods

Species morphospace and sequencing data

We compared the floral phenotypic similarities of the six species sampled (see details in chapter 4), using 10 floral measurements available from the literature (Perret *et al.*, 2007; Serrano-Serrano *et al.*, 2015). Data was log-transformed and used in a Principal Component Analysis (PCA) in the R package function *prcomp* in order to distinguish relevant traits producing the differentiation between pollination syndromes. RNA extraction, library construction and sequencing, transcriptome assembly, quality and orthologous search is described in chapter 3. Here, we used all the one-to-one orthologous groups (OG) with the associated information of expression (raw counts and gene length), and the sequence alignments. All the comparisons conducted here recall our experimental design (see Figure 1), which combines the species with the same pollinator type, allowing us to compare the potentially convergent patterns. Thus, all results are presented and discussed in light of the contrast between two phenotypes: bee and hummingbird specific patterns.

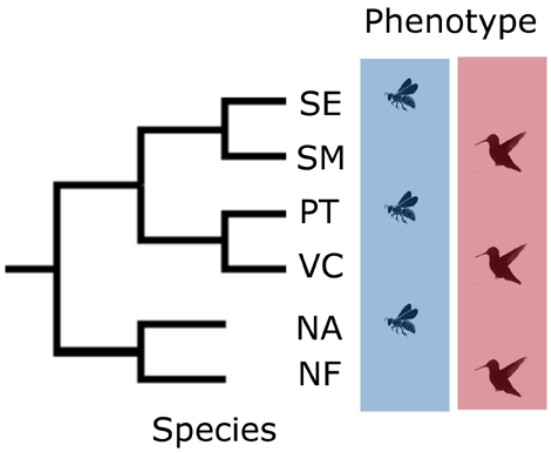


Figure 1. Experimental design for the evaluation of concerted patterns. At the gene expression level we analyzed the different developmental stages separately, but biological replicates are combined or accounted in the statistical models. At the molecular signatures, a single gene sequence represent all stages and replicates within a species. With this framework, all references in the text to "bee-pollinated species" contain the information from SE, PT, and NA species, while all references to "hummingbird-pollinated species" contain the information from SM, VC, and NF. Patterns referred to "global" contain the information from the full tree.

Expression analyses

Gene expression normalization

The total abundance matrix representing the expression of each OG in every species and condition was estimated in chapter 3. However, these values require normalization, and there are three sources of variation in the abundance matrix that make the raw counts from different replicates and experimental conditions not directly comparable (Robinson & Oshlack, 2010). First, differences between

libraries arise from each library size (sequence depth of each experiment) with larger library sizes producing higher total counts. Second, variation within libraries due to the differences in gene size, longer genes are expected to have larger number of reads mapping. Third, due to the RNA composition of each sample, which refers to the amount of highly expressed genes within a library that modifies the amount of reads that can align to the lowly expressed genes. Indeed, the distribution of gene expression counts is skewed with few genes having very large values, and the variances around these counts increase with expression level (Zwiener *et al.*, 2014).

These sources of variation were circumvented by data transformation. Here, we first accounted for differences in gene size between species assemblies using the gene length from each species (rpkm function in edgeR package, Robinson *et al.*, 2010). Expression values were then normalized to consider the variability between genes and conditions (named the mean-variance relationship), using the *voom* function in the limma R package (Ritchie *et al.*, 2015). We applied the *cyclic-loess* normalization which is robust to unbalanced differential expression and incorporate probe-wise weights (Law *et al.*, 2014). This approach has been shown to better control type I errors and False Discovery Rate (FDR) even with a low number of samples (Soneson & Delorenzi, 2013). This procedure incorporated a design matrix to identify replicated samples within conditions, batch effects, and phylogenetic relationships (by specifying which pairs of species are more related). These transformations gave us the final normalized gene expression matrix.

Differential gene expression

A normalized gene expression is the initial data set to find gene expression differences between conditions. We checked the quality and differentiation between RNA-Seq libraries and conditions through the multidimensional scale plots (MDS plot), using the unsupervised clustering method implemented in the plotMDS function in the limma R package (Ritchie *et al.*, 2015). We investigated the differentially expressed genes (DEG) between the pollination syndromes by fitting a gene-wise linear model. This model contrasts the pollinator-specific expression within each developmental time. We performed the multiple t-statistics with the *nested* method that provides greater weights to genes that are significant in more than one contrast, to enhance those genes that have responses at multiple developmental times. A minimum log-fold change of 2 units and a minimum p-value of 0.05 were considered to estimate the significantly DEG. We retrieved the Gene Ontologies (GO, http://geneontology.org/) for all annotated DEG, and GO enrichments analyses were performed using Uniprot IDs and *A. thaliana* as reference. Indeed, the identification of over-represented GO is plausible if many genes are associated to the same metabolic route or function. However, if there is a rather simple genetic control of floral traits, with few genes producing major phenotypic differences, the probability of finding enriched GO is low. Finally, to help the understanding of the patterns of differential expression we classified the genes by developmental stage, phenotype-specific expression, and their functional annotation.

Evolutionary rates and selection signatures

Evolutionary rates

We performed multiple sequence alignments (MSA) of the peptide sequences for each one-to-one OG using MAFFT (Katoh & Standley, 2013). The mRNA sequences were translated into codon alignments using Pal2Nal (Suyama *et al.*, 2006). We constrained the tree topologies to follow the species-tree and estimated the branch lengths and parameters of the model of substitution (General Time Reversible, GTR) with *PhyML* (Guindon *et al.*, 2009). We used the total branch length as a proxy of the total evolutionary rate, log-transformed these measurements, and first categorized the genes as rapidly

and slowly evolving (2.5% upper and lower percentiles, respectively). These genes have global rate patterns, however the genes with phenotypic specific rates patterns (i.e. rapid in all hummingbird- or bee-pollinated species) were also targeted. For this, we measured a ratio for each gene as the mean evolutionary rate for hummingbird-pollinated species divided by the mean evolutionary rate for bee-pollinated species and selected the genes with rate ratios displaying at least a 2-fold change. Positive values of the ratio indicate rapid evolution in all hummingbird-pollinated species, while negative indicate in bee-pollinated species respectively.

Selection signatures

We investigated the selection signatures in the set of one-to-one OG. We used the program codeml from PAML (Yang, 2007) to calculate the dN/dS ratio (nonsynonymous dN, synonymous dS substitutions). Under the null model (H_0 = M2a_rel) we expect codon positions to evolve in three different classes: neutrally evolving sites ($\omega \approx 1$), sites evolving under purifying selection ($\omega < 1$), and sites evolving under diversifying selection ($\omega > 1$). This null model assumes that all branches on the tree have the same ω values, while the alternative model (H_1 = CladeC) allows clade-specific ω values. We defined phenotype-specific sets of branches: ω_3 and ω_4 containing all bee-pollinated and all hummingbird-pollinated species, respectively (Fig. 2). The H_0 and H_1 models were compared using Likelihood Ratio Tests (LRTs) p-values after correcting for FDR in multiple testing (q-values estimated in R package). The statistical power of the LRTs depends in large part on the number of sequences present in the phylogenetic tree. We therefore increase the size of each OG by enriching them with additional outgroup sequences for a minimum set of 7 species, the six Gesnerioideae plus an outgroup (Table 1). We performed a second orthologous search with OrthoMCL to identify one-to-one OG set between the OG groups defined four our six species and at least one outgroup sequence. The LRTs identified OG where the alternative model was significantly better (q-value < 0.01), and these were classified in three categories: OG with signatures of purifying selection in any of the phenotypes (ω_3 or $\omega_4 < 1$), OG with signatures of divergent selection in any of the phenotypes (ω_3 or $\omega_4 > 1$), OG with signatures of divergent selection in both of the phenotypes (ω_3 and $\omega_4 > 1$).

Results

Floral morphological space and pollination syndromes

The six selected species showed convergent traits regarding the pollination interactions. We illustrated the differences in floral morphologies and phylogenetic relationships between them in Fig. 3. Bee-pollinated species have whitish open corollas with a landing platform, while hummingbird-pollinated species have red or pink tubular corollas. A large portion of the variance in floral morphospace is explained by the first and second axes of the PCA analysis (24.91%, 16.08% respectively, see loadings in Table 2). PC1 loadings are associated with the shape of the corolla and the reproductive organs, specifically with the length from the anther/stigma to the ovary (LAN and LST, which determine for exserted anthers for the hummingbird pollinated flowers in the positive side of the plot), and the horizontal diameter of the flower opening (DOH, for very narrow corolla mouth for the hummingbird-pollinated flowers, in contrast with more open corollas of the bee-pollinated ones). PC2 loadings are associated with the total length of the flower (LTU in fig. 3), with smaller flowers in the positive values (*V. calcarata*, and *S. eumorpha*). These results supported a distinct floral morphospace for the two pollination syndromes, particularly determined by floral shape and size (see specific trait values in fig. 4).

Site class	0: purifying		1 : Neutral		2 : divergent	
Model	ω (dN/dS)	Proportion	ω (dN/dS)	Proportion	ω (dN/dS)	Proportion
Clade model C	$0 < \omega_0 < 1$	p_0	$\omega_1 = 1$	p_1	$\omega_2, \omega_3, \omega_4 > 0$	$1 - p_0 - p_1$
M2a_rel	$0 < \omega_0 < 1$	p_0	$\omega_1 = 1$	p_1	$\omega_2 (=\omega_3 = \omega_4) > 0$	$1 - p_0 - p_1$

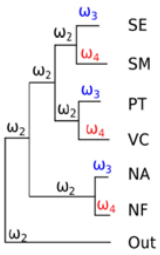


Figure 2. Two model comparison for the selection tests (modified from Weadick & Chang, 2012). Clade model C contains 6 parameters (p_0,p_1,w_0,w_2,w_3,w_4), and the null model M2a-rel contains 4 parameters (p_0,p_1,w_0,w_2=w_3=w_4). Blue and red branches show the bee- and hummingbird-pollinated species, respectively.

Table 1. Outgroup sequences included in the selection tests.

Species	Database*	Type of data	Number of sequences
Solanum lycopersicum	plantGDB	mRNA	56'845
Aquilegia formosa x Aquilegia pubescens	plantGDB	mRNA	19'615
Petunia axillaris	plantGDB	mRNA	25'774
Petunia integrifolia	plantGDB	mRNA	27'341
Ipomoea nil	plantGDB	mRNA	22'946
Silene dioica	plantGDB	mRNA	20'046
Silene latifolia	plantGDB	mRNA	39'700
Mimulus guttatus	Phytozome	Proteins and CDS	28'140
Arabidopsis Thaliana TAIR10	Phytozome	Proteins and CDS	27'416

*All sequences were all downloaded on June 3rd, 2015. Phytozome V10.2, and plantGDB Release 187 (2011-12-15).

Comparative gene expression in the Gesneriaceae species

The overview of the expression patterns between the six Gesneriaceae species suggested that normalized expression values are clustered by developmental stage and species, with two strongly separated clusters of vegetative and floral libraries (green and pink boxes in Fig. 5). Within floral libraries the small buds (B1) clustered independently from the middle and adult flowers (B2 and FL). However, these two stages were separated by taxonomy (with separated *Nematanthus* genus and Ligeriinae lineage clusters). In most cases biological replicates clustered together (with exception of B1 stage for SM and VC, and VG stage for PT, see those cases as gray circles in Fig. 5). In addition, libraries clustered by taxonomic relationships between pairs of more related species (gray rectangles in Fig. 5).

We performed a principal component analyses to further dissect the variation in gene expression patterns (Fig. 6). The largest proportion of the variance is explained by developmental stages (PC1 = 16.79% variance) and taxonomic relationships (PC2 = 14.20% variance). This pattern agrees with the previous overall clustering of libraries, with the main separation of vegetative and floral samples. PC3 and PC4 axes explained jointly an additional 19.13% of the variation, with the three pairs of closely related species clustering together. Finally, the next two axes showed the separation of the pollination syndromes (Fig. 6C). This separation explains 13.81% of the variance, with PC5 discriminating *S. magnifica* and *S. eumorpha*, and PC6 the two remaining species pairs.

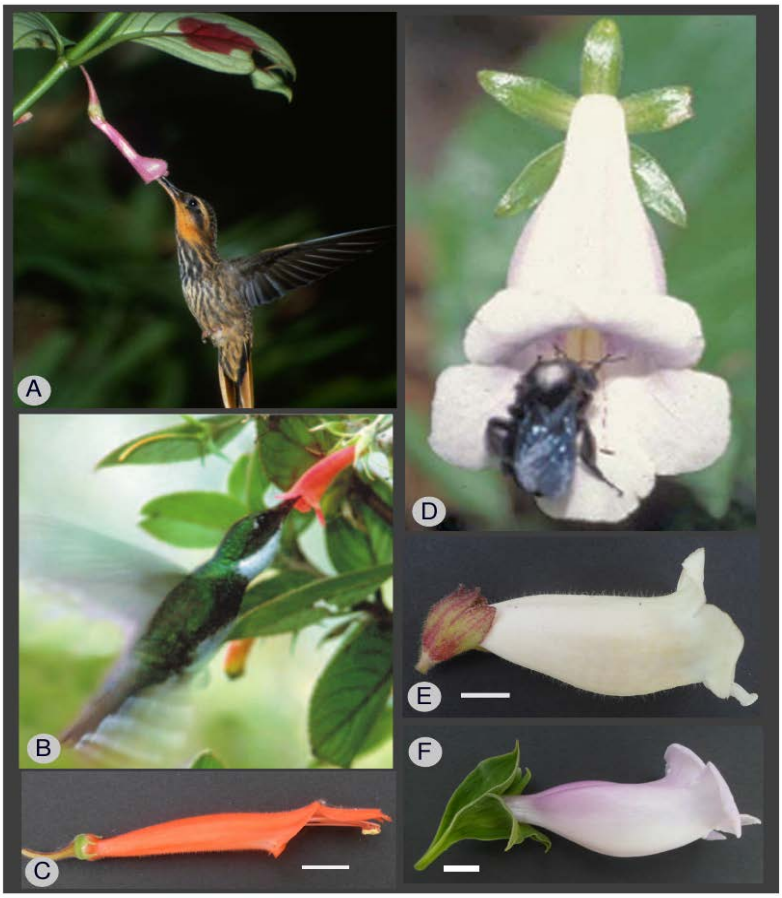
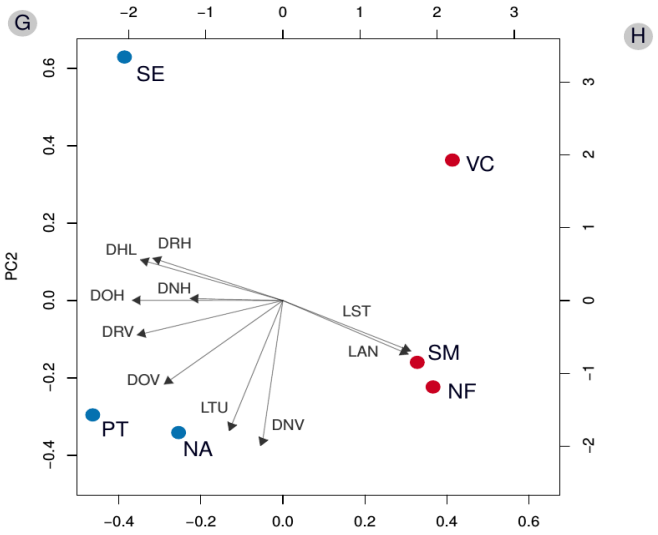
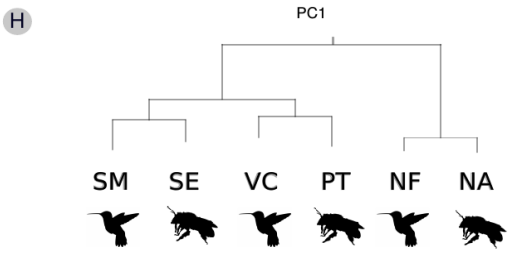


Figure 3. Plant-pollinator interactions, floral morphology, phenotypic space and phylogenetic relationships of the six Gesneriaceae species. A) *N. fritschii* (NF) visited by *Ramphodon naevius* (photo by Silvana Buzato). B) *V. calcarata* (VC) visited by *Leucochloris albicolis* (photo from SanMartin-Gajardo & Sazima, 2005). C) *S. magnifica* flower (SM). D) *S. eumorpha* (SE) visited by *Bombus morio* (photo from SanMartin-Gajardo & Sazima, 2004). E) *N. albus* flower (NA). F) *P. tenuiflora* flower (PT). Scale bars: 1 cm in C, E, and F. G) Principal component analysis for 10 floral traits, colors correspond to the pollination systems (blue for bee-, red for hummingbird.). H) Phylogenetic relationships between the species, subset from chapter 2, figure 1.





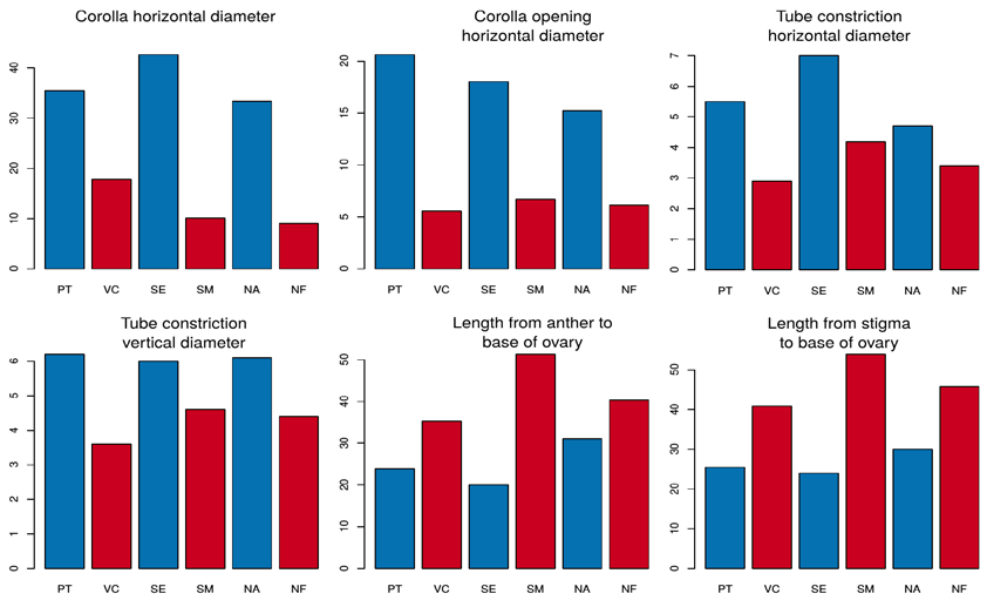


Figure 4. Floral traits differences between pollination syndromes. Colors represent bee- (blue) and hummingbird-pollinated species (red). Acronyms correspond to the species names described in the legend of fig. 3. The scale of the vertical axis correspond to millimeters.

Table 2. PCA loadings for floral morphospace.

Trait	PC1	PC2
Corolla horizontal diameter (DHL)	-0.3575	0.1933
Dorsal tube length (LTU)	-0.1467	-0.5700
Horizontal diameter of the corolla opening (DOH)	-0.4010	-0.0042
Corolla vertical diameter at the corolla opening (DOV)	-0.3255	-0.3498
Horizontal diameter of the tube constriction anterior to the nectar chamber (DRH)	-0.3646	0.1020
Vertical diameter of the tube constriction anterior to the nectar chamber (DRV)	-0.3834	-0.1341
Horizontal diameter of the nectar chamber (DNH)	-0.2419	0.0621
Vertical diameter of the nectar chamber (DNV)	-0.0363	-0.6171
Length from the anther base to the ovary (LAN)	0.3320	-0.2726
Length from the stigma base to the ovary (LST)	0.3660	-0.1672

Differential gene expression between phenotypes

We found 3.9% of the total number of genes (288 out of 7287) that are differentially expressed in all the pairs of species with contrasting pollinator types. Many of those DEG are shared across developmental stages, and 38 (13.2%) of them are constitutive, or DEG in all pairs of species at all developmental times (Fig. 7A). The number of DEG restricted to a single developmental stage increased flower maturation (B1 = 18 DEG, B2= 24 DEG, and FL= 42). For all stages, we found a larger proportion of up-regulated DEG in bee-pollinated than hummingbird-pollinated species (Fig. 7B) and this patterns holds if we remove the vegetative libraries (equal number of DEG, data not shown). We found no differences in the absolute log-fold change (lfc) between floral and vegetative libraries (all t-tests > 0.05, Fig. 7C). However, we observed that DEG shared across all stages showed a slight increase in lfc through time, with late stages showing larger lfc differences in expression (Fig. 7C).

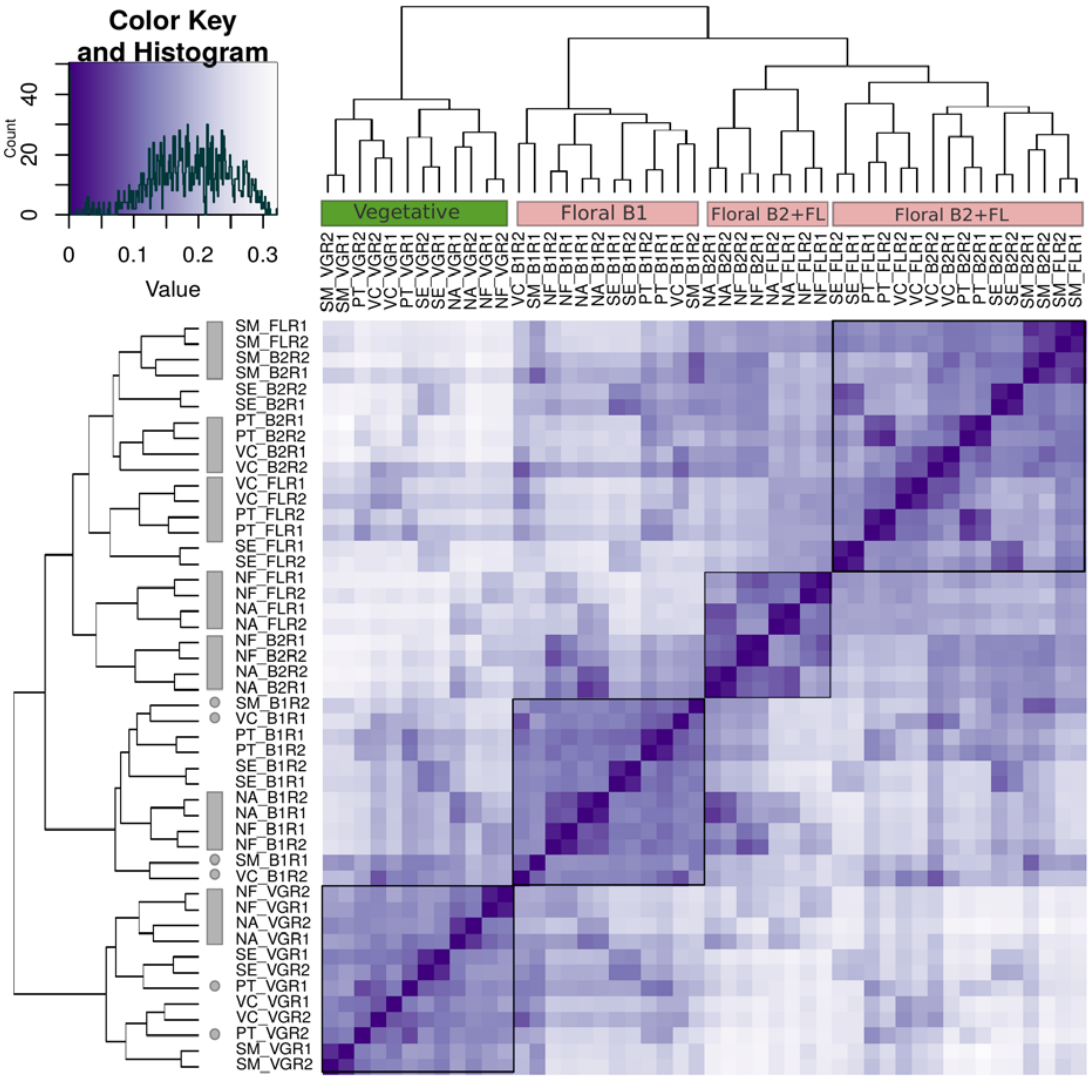


Figure 5. Heatmap of library-to-library distances using the whole matrix of normalized expression data. Green and pink boxes on top indicate vegetative and floral libraries, respectively. Gray boxes on the left show libraries from closely related species. Gray circles show biological replicates that do not cluster together.

More than half of the DEG has no predicted functional annotation (57%), and only 124 transcripts were associated with known proteins (Table 3, and full set in supplementary material S1 on line). Among the annotated transcripts 29 (23%) have regulatory functions, with 18 bee-, and 11 hummingbird up-regulated. Moreover, distinct transcription factor (TF) families were found differentially expressed between phenotypes, with AP(1), SCL(32), AGL(9,15), AMS, and DOF(53) up-regulated uniquely in bee-pollinated species. In contrast, TF families bHLH(75), bZIP(61), RF2b, ATHB(5), AHL(10), WOX(2), and GRF(9) were up-regulated in hummingbird-pollinated species. Finally, TF families EF, MYB, MADS, NAC were present in both phenotypes.

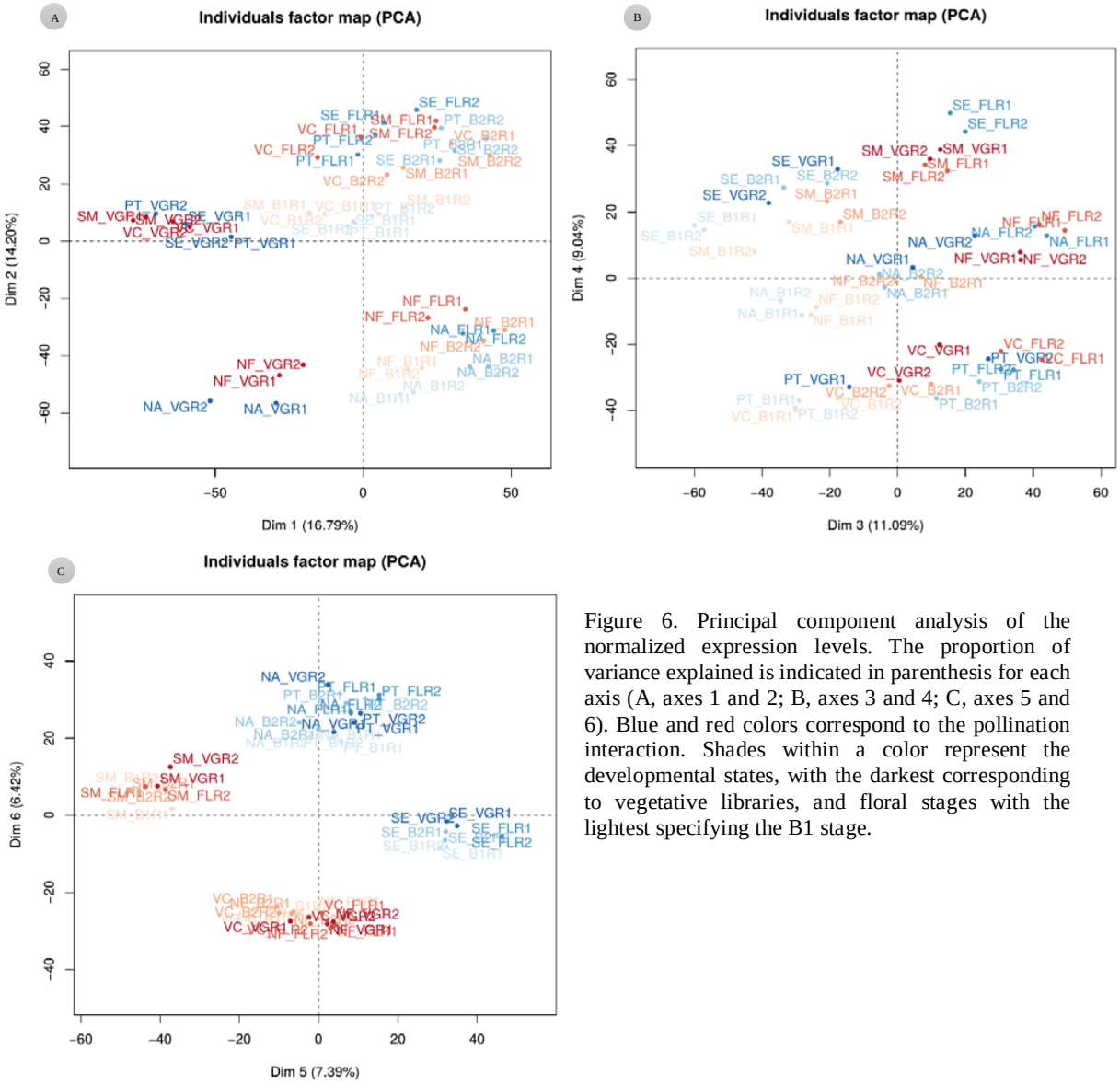


Figure 6. Principal component analysis of the normalized expression levels. The proportion of variance explained is indicated in parenthesis for each axis (A, axes 1 and 2; B, axes 3 and 4; C, axes 5 and 6). Blue and red colors correspond to the pollination interaction. Shades within a color represent the developmental states, with the darkest corresponding to vegetative libraries, and floral stages with the lightest specifying the B1 stage.

We also found several genes differentially expressed between the two phenotypes. A total of 18 OGs showed constitutive DEG up-regulation in hummingbird-pollinated species, and this set of genes included the DTX41 Detoxification 41 protein (also called TT12 Transparent Testa). Similarly, 20 OGs showed constitutive DEG up-regulation in bee-pollinated species, and this set contained a TF acting on indolic glucosinolates biosynthesis, and potentially in plant defenses. Early and middle buds (B1 and B2 stages) had no convergent functions between phenotypes, except for a series of very generalist proteins associated with photosynthetic functions. In contrast, adult flowers and vegetative material in both phenotypes showed potentially relevant functions for the pollination syndrome traits. First, up-regulation in adult flowers from hummingbird-pollinated species showed multiple regulatory TF (Wuschel homeobox, AT-hook motif, and MADS-box types), and proteins causing loosening and extension of plant cell walls (expansin-15). Adult flowers from bee-pollinated species showed up-regulation of antifungal aromatic compounds, geranyl and farnesyl biosynthesis (GGPPS), TF associated with longevity regulation (JUNGBRUNNEN 1), flavonoid 3'monooxygenase (F3PH), which drives the anthocyanin pathway towards more cyanidin-derived compounds (non-red pigments).

Finally, DEG for hummingbird-pollinated species at vegetative stage included reductase proteins involved in lignin biosynthesis, and TF for cell expansion, while vegetative DEG in bee-pollinated species included proteins for pollen tube growth (BABL), modulation of rotational polarity and anisotropic cell expansion (WVD2), early floral meristem identity and anthocyanin production (DFR) proteins.

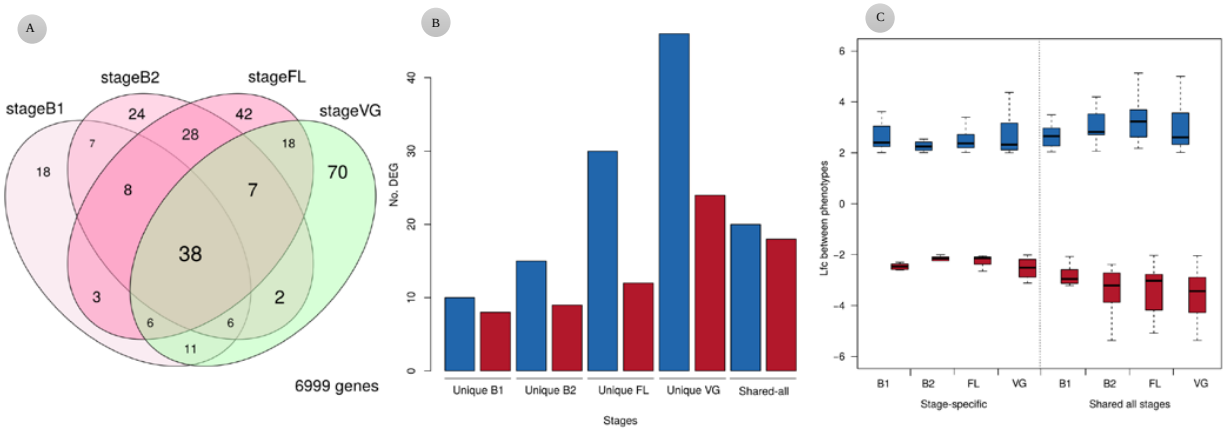


Figure 7. A) Venn diagram of the number of DEG between pollinator phenotypes across developmental stages. B) Number of genes up-regulated in bee-pollinated (blue) and hummingbird-pollinated species (red). Bars show only the stage-specific DEG (labelled unique) and those shared across all stages. C) Log-fold change between phenotypes for DEG at stage-specific or shared, colors as in section B.

Table 3. Functional categories for DEG up-regulated in bee- and hummingbird-pollinated species.

Category	**Bee-specific functions**	**Hummingbird-specific functions**
Constitutive (DEG in all developmental stages) N = 38	Translation, lipid transport, *defense response*, ethylene signaling, cell wall organization	Translation, response to stress, glycosylation, ion transport, *flavonoid process*
Unique to flowers (DEG not in vegetative, but unique (pink boxes below) or shared (46 genes) between developmental stages) N = 130	Defense response, lipid metabolism, ethylene signaling, flavonoid biosynthesis (see additional categories for unique stage DEG)	Carbohydrate metabolism, chaperone-mediated folding, cytoplasm (see additional categories for unique stage DEG)
Unique to vegetative N = 70	Regulation of transcription, glucosinolate biosynthetic process, photosynthesis, GA and auxin signaling, *floral meristem*, defense response, DNA repair, *flavonoid biosynthesis*	Regulation of transcription, response to stress, developmental process, *lignan biosynthesis*, *transmembrane transport*
Unique to B1 stage N = 18	Regulation of transcription, polysaccharide metabolism, defense response, abscisic acid signaling, histone acetylation	Regulation of transcription, defense to fungal infection, cellular detoxification, cytoskeleton
Unique to B2 N = 24	Cellular response to DNA damage, glycerol transport, fatty acid biosynthesis, Notch signaling	Auxin homeostasis, photosynthesis and chloroplast, folic acid biosynthesis
Unique to FL N = 42	Regulation of transcription, *anthocyanin* and *carotenoid biosynthesis*, lipid metabolism, oxidation process, protein folding	Regulation of transcription, multicellular development, *cell wall organization*, cell redox homeostasis

Evolutionary rate analyses and selection signatures

Evolutionary rate analyses

We used the distribution of total branch lengths to identify the OGs present in the upper and lower 2.5% percentiles. We found 442 OGs representing the rapidly and slowly evolving genes in the set of species (Fig. 9). Among the rapidly evolving genes 14 corresponded to poor quality alignments (manually examined), and only 50% of all rapidly evolving were functionally annotated (Supp. Mat. S1). The GO enrichment analysis of those rapidly evolving genes suggested that meristem development (tissue and organ development), anatomical morphogenesis, macromolecule localization and metabolic processes ontologies were enriched (at >2-fold change, p.value <0.05). In contrast, the slowly evolving genes were enriched for DNA-templated transcription and elongation, vesicle-mediated transport, and macromolecular subunit organization (Supp. Mat. S1 on line).

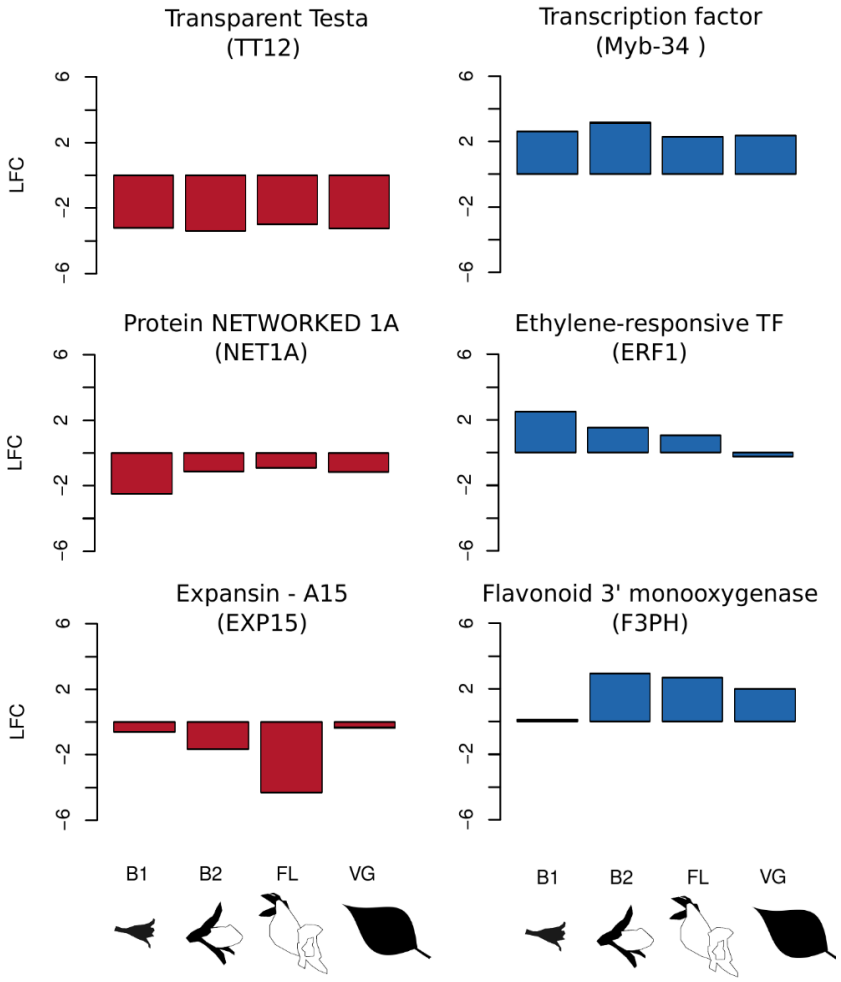


Figure 8. Examples of gene expression data between pollination types. The bars correspond to the log-fold expression relative to each other pollination type (red: hummingbird-pollinated species, blue: bee-pollinated species). Pictograms represent the developmental stages for each expression bar.

The differences in branch length between bee- and hummingbird-pollinated OGs for the three pairs of species allowed the identification of phenotype-specific rapidly evolving genes (see example in fig. 9 C and D). The proportion of those rapidly evolving genes is low and did not differ between phenotypes (1.06% and 1.32% for hummingbird- and bee-pollinated species, respectively, see fig. 9B). The annotation of those genes for both phenotypes are presented in Supp. Mat. S1 on line. Genes with rapid evolution in hummingbird-pollinated species were enriched for the metabolic process of organic substance (p-value = $2.89x10^{-2}$) with proteins related to carbohydrate metabolic process proteins (β-galactosidase 17, β-glucosidase 47, and malate dehydrogenase), protein folding and transport (NACA2, CYP23), and circadian regulation (Reveille 6 and 8 proteins). Genes in bee-pollinated species with rapid evolution were enriched for anatomical structure development (p-value = $2.17x10^{-2}$), with multiple cyclin proteins (A1-1, H1-1), floral meristem regulator ULTRAPETALA-1, transcription factors (Myb-APL and PCF7).

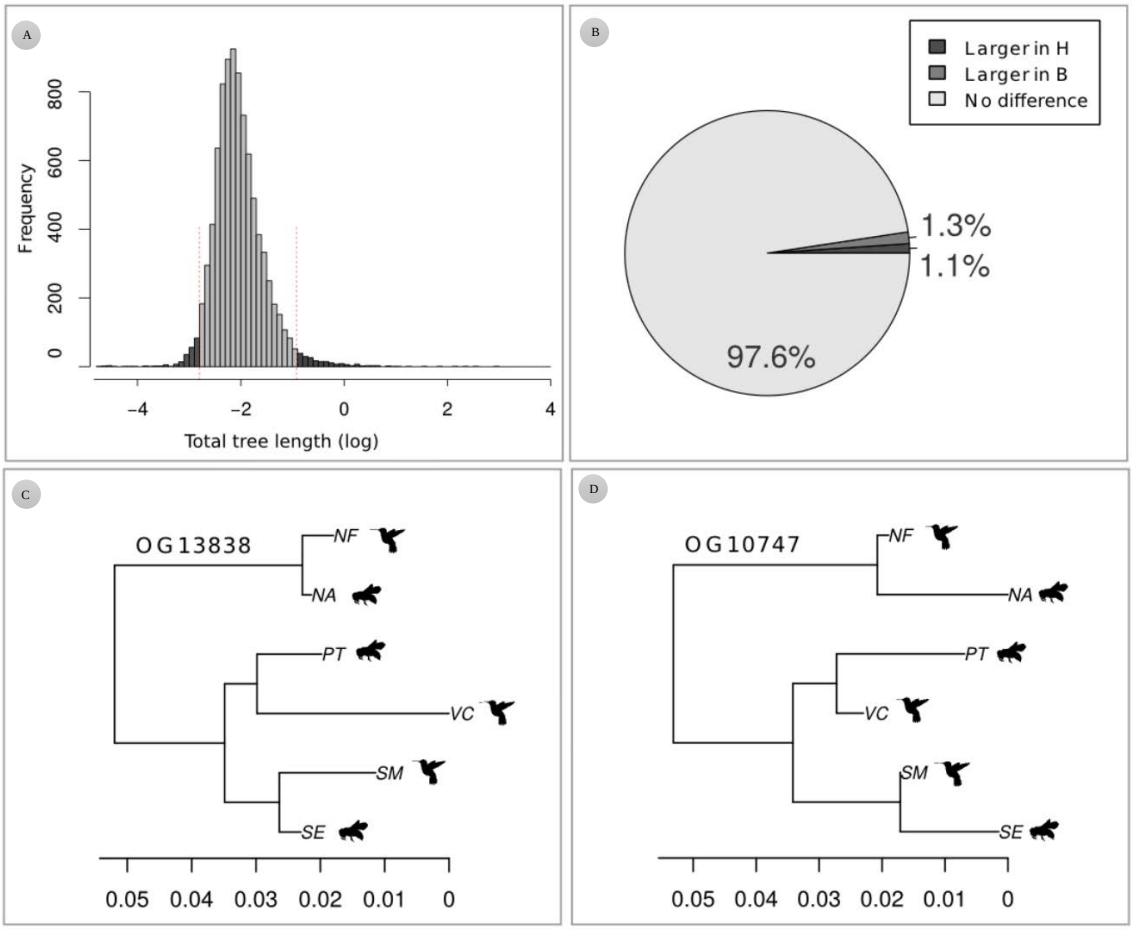


Figure 9. A) Distribution of total tree length for 8848 OG, darker bars in histogram correspond to the upper and lower 2.5%. B) Proportion of rapidly evolving genes in hummingbird- (H index) and bee-pollinated species (B index). C and D) Examples of genes with rapid evolution in hummingbird-pollinated and bee-pollinated species, respectively.

Purifying selection signatures

The total number of genes identified with footprints of selection was 1067 (14.02% of 7612 OG). The number of genes that showed signatures of purifying selection was 350 (4.59%). Among those **148** showed at least a 2-fold stronger purifying selection in hummingbird-pollinated species (median omega

value of 0.073), and **202** genes where bee-pollinated species showed at least a 2-fold stronger purifying selection (median omega value of 0.056). A total of 717 genes showed signatures of positive selection (11.00%, all positive selection in Fig. 10). Those genes were classified in three categories: genes with positive selection only hummingbird-pollinated species (A, n=271), genes with positive selection only in bee-pollinated species (B, n=230), and genes with positive selection in both phenotypes (C and D, n=216). Among the latter category, 120 have at least a 2-fold larger omega in hummingbird-pollinated species, and 96 have the same pattern in bee-pollinated species.

The annotation of genes under purifying selection in hummingbird-pollinated species was enriched for glucan biosynthetic process (p-value $3.76x10^{-2}$), organic hydroxyl biosynthetic process (p-value $2.50x10^{-2}$), monocarboxylic acid metabolic process (p-value $4.70x10^{-4}$), and organic substance catabolic process (p-value $3.63x10^{-5}$). Among these genes, we found functional categories such as carotenoid process (LCYE, lycopene cyclase), modification of cell wall structure and loosening (SBT1.7, EP1), and pollen tube elongation and polar growth (VP52A). The genes under purifying selection bee-pollinated species presented more general enrichment categories, with single-organism metabolic process (p-value $5.20x10^{-3}$), cellular metabolic process (p-value $1.72x10^{-2}$), and organic substance metabolic process (p-value $2.82x10^{-2}$). Potentially relevant functions included actin-based cell morphogenesis (NAP1), auxin homeostasis (TCP15), calreticulin-3 acting on nectaries and anthers (CALR3), and flavin-flavonoid process (NUD23, F3PH).

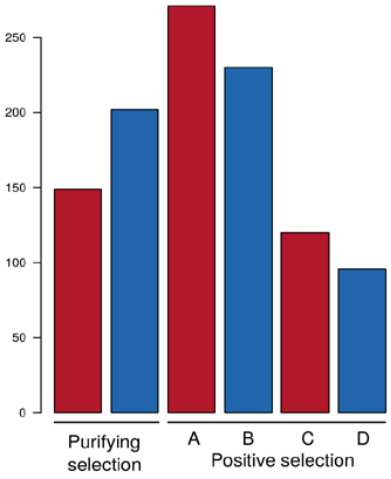


Figure 10. Number of statistically significant genes with purifying and positive selection. Blue and red bars indicate the genes where bee- and hummingbird-pollinated showed signatures of selection, respectively. A and B bars correspond to genes where only one of the phenotypes show positive selection. C and D bars are genes where both phenotypes showed positive selection, but one (red or blue, C or D) has larger omega value.

Positive selection signatures

The genes that showed evidence of positive selection were enriched in hummingbird-pollinated species for categories such as cofactor biosynthetic process (p-value $3.20x10^{-3}$), protein import (p-value $5.10x10^{-3}$), DNA recombination (p-value $3.05x10^{-2}$), ncRNA processing (p-value $1.22x10^{-2}$), and histone modification (p-value $4.45x10^{-2}$). Within these categories multiple transcription factors were found to be under positive selection (AMS, RAX2, MADS4, TADA2, bHLH93, bHLH140, HSFA2, HFB2B, RVE8, TRAB1, and WRK26). Among those, the TTL1 a transcriptional regulator of flavonoid accumulation (ZWIP6). Categories enriched for positive selection in bee-pollinated species included nuclear division (p-value $2.51x10^{-2}$), macromolecular complex organization (p-value $1.95x10^{-2}$), nucleobase-containing compounds (p-value $1.90x10^{-2}$). Multiple genes were involved in potentially relevant functions such as cell morphogenesis, with cell elongation and orientation (FKB42), lateral organ polarity (KAN3, ARC2A), reorientation of microtubules (KTNA1), anisotropic cell expansion (WDL1), polarized cell growth (SEC6, IDD14, LBD13) proteins, and glucosinolate biosynthesis (BASS2, GSOX5) or essential plant oils (Isoeugenol IGS1).

Common genes between DEG, evolutionary rates and selection patterns

Two OG (OG13954 and OG6634) were found to be common to the three analyses that we performed above (DEG, evolutionary rates and selection signature). None of these genes have a predicted annotation, but both are up-regulated in hummingbird-pollinated species and showed larger substitution rates and strong evidence for positive selection in the same phenotype. Only 7% of the DEG (20 genes) were found in the overall rapidly evolving gene category, with functions related to protein folding, transcription regulation, floral whorl development (AGL9, AP1, MADS-SVP), and flavonoid metabolic process (SOT5). Most of these genes were up-regulated in bee-pollinated species. We further contrasted the phenotype-specific rapidly evolving genes and all the DEG, and found only four genes that were in common between the two analyses. These include two rapidly evolving genes in hummingbird-pollinated species (OG12614 of unknown function, and OG15759 a predicted NACA2 protein, both up-regulated in bee-pollinated flowers at all developmental stages), and two rapidly evolving genes in bee-pollinated species (OG36280 a flavonoid sulfotransferase SOT5 and OG5375 a trichome birefringence-like protein TBL3 in charge of cell wall deposition, both up-regulated in bee-pollinated species).

For genes identified for both the signatures of selection and DEG (36 OG), there is no clear pattern between the positively selected and the phenotype-specific expression. Within the positively selected genes in bee-pollinated, 11 are up-regulated in bee- and 6 in hummingbird-pollinated species. Similarly, the positively selected genes in hummingbird-pollinated species were up-regulated in either bee- (6 genes) or hummingbird-pollinated (3 genes) species. Only two cases of overlapping signals concerned relevant functions: i) the up-regulation of TT12 (Transparent Testa 12, OG30097) in all developmental stages, and the positive selection of the ZWIP6 (OG35706) both in hummingbird-pollinated species, ii) the up-regulation of F3PH in later stages of floral development (OG11385), and the purifying selection of the same OG, both in bee-pollinated flowers.

Discussion

The degree at which nature find similar solutions for similar environmental pressures is astonishing, and the genetic mechanisms for such evolutionary outcomes are suggested to be highly conserved (McGhee, 2011; Martin & Orgogozo, 2013). Floral phenotypes are highly similar evolutionary solutions that facilitate the attraction and reward of pollinators, to ensure plant reproduction. These phenotypes involve morphological changes potentially driven by concerted genetic mechanisms. However, the study of these mechanisms in plants, using a macro-evolutionary approach, is still in early stages (Bartlett & Specht, 2010; Zhang *et al.*, 2012; Hileman, 2014). Here, we examine the contribution of concerted changes in gene expression (regulatory changes) and coding sequence (structural changes) to the evolution of convergent pollination syndromes in flowers of the Gesneriaceae family

Concerted responses at regulatory and structural level

Phenotypic differences between organisms are thought to be controlled by two processes: divergences in the protein sequence and gene expression patterns. Both of them have shown to evolve highly correlated in mammals and insects (Hunt *et al.*, 2012; Warnefors & Kaessmann, 2013). Explorations in plants are limited, and mostly evaluated within single gene families (Bartlett & Specht, 2010). Our results suggest that these two mechanisms may have worked independently during the parallel evolution of floral morphologies, with a larger contribution from concerted sequence divergences, than concerted gene expression changes. We report and discuss here the specific patterns identified. First, we found a low amount of concerted regulatory changes during the floral evolution of the Gesneriaceae

species. The overall characterization of the transcriptomes of the six species indicated that the expression profiles are highly similar between all sampled species, with a clustering based on the developmental stages (Fig. 5). The tendency of global patterns of gene expression to cluster by tissues or developmental times is coherent with the findings in comparative transcriptomes among mammals (Brawand *et al.*, 2011), suggesting that differences in gene expression could vary within specific tissue, organ or developmental stage programs (Pankey *et al.*, 2014). Only a small proportion of variance in gene expression is due to the phenotypic differences associated with pollinator type (Fig. 6C). This general pattern coincides with the low percentage of concerted DEG between pollinator transitions (3.9% of the transcript dataset). Studies in comparative transcriptomic have the goal to identify differences in gene expression that may explain changes in phenotypes between species, populations or tissues. However, there is the difficulty of distinguishing whether the observed differences in gene expression, between or within a species, are the result of adaptive differentiation, or rather correspond to neutral divergences (Khaitovich *et al.*, 2004). This controversy has motivated the development of alternative ways to model gene expression (Harrison *et al.*, 2012; Rohlfs *et al.*, 2014), and the implementation of robust experimental designs during the characterization of transcriptomic responses (DeBiasse & Kelly, 2016). The timescale (between 4.7- 12.5 Mya for species divergences), multiple replicates of the pollinator shifts, and the incorporation of the species relationships in the models for differential expression used here, should maximize the possibility that the observed transcriptomic differences are due to adaptive and coordinated responses in targeted pathways, more than possible neutral evolution.

A large proportion of the concerted DEG are up-regulated in bees (Fig. 7B), suggesting a higher level of concerted expression patterns for this phenotype. This result could be due to the fact that bee-pollination is the ancestral state for the subfamily, and similar mechanisms operate to produce bee-pollinated flowers. In contrast, the repeated evolution of hummingbird pollination may have been produced by slightly alternative genetic mechanisms, which would reduce the extent of concerted gene expressions in hummingbird-pollinated plants. This finding agrees with a study of the evolutionary trends in floral transcriptomes, which has shown that basal species (though at a large scale in the angiosperm evolution) have strong similarities in their expression profiles, while derived species show organ-specific transcriptional programs, which may correlate with the evolution of morphologically distinct flowers (Yoo *et al.*, 2010). A complementary exploration, allowed by the current data, could include the evaluation of gene expression differences on each pair of related species, providing us information on the specific expression profiles, and whether there is an overlap at functional pathways between the evolutionary shifts.

The amount of concerted DEG increased with developmental time, with adult flowers showing two times more concerted responses than early stages, indicating a potential contribution of the later developmental stages to the final floral morphology. Unexpectedly, vegetative tissues showed the largest amount of concerted DEG (Fig. 7B). These common genes, identified in vegetative tissues, may reflect concerted responses in the entire plant morphology, such as defense and stress responses, lignin biosynthesis, and photosynthesis (see Table 3), potentially involved in additional ecological or physiological shared conditions between the species investigated .

Secondly, the contribution of structural changes, determined by the proportion of OG with footprints of selection at specific pollination syndromes, was higher than the regulatory changes (14.02% versus DEG proportion 3.90%). A similar amount of genes have been found with signatures of positive selection between two closely related species of primroses (13% of OG between *Primula poissonii* and *P. wilsonii*, Zhang *et al.*, 2013). This pattern most likely reflects a stronger influence of sequence than expression divergence, in the genetic mechanisms for floral convergences. Specific patterns of selection showed that

4.59% and 9.42% of the OGs were evolving under purifying and positive selection, respectively. The amount of OG under purifying selection was larger for bee-pollinated than hummingbird-pollinated species (Fig. 10), suggesting that the functionality of producing a bee-flower is rather maintained, preventing degeneration or somehow reflecting less global variation in this phenotype. On the contrary, hummingbird-pollinated species showed a larger number of positively selected genes than bee-pollinated, indicating that sequence divergence is likely involved in the evolutionary changes leading to hummingbird pollinated flowers. The presence of positive selection has been suggested to underlie biochemical diversification of plant defenses (Benderoth *et al.*, 2006), and the convergent evolution of phenotypic adaptations to aquatic environments in mammals (Foote *et al.*, 2015). Moreover, this pattern agrees with the hypothesis that positive selection is a way to modify cis-regulatory sequences, and the overall molecular machinery for the origination of a new phenotype, without the strong constraint of pleiotropic effects (Wray, 2007).

Recent studies found evidence that rapidly diverging genes, but not necessarily under positive selection, tend to have larger variance in expression levels, and potentially a contribution in the adaptive responses of conifers (Hodgins *et al.*, 2016). In our transcriptomes data, several of the rapidly evolving genes also displayed differential expression between the two phenotypes, especially for categories such as floral whorl development and flavonoid metabolic process, two functions highly involved in the convergent floral phenotypes examined. Despite these few examples, there is overall very little overlap between the rapidly evolving genes or those with signatures of selection and the DEG. This suggests that the differences in gene expression between phenotypes may not be due to fixed genetic changes in protein coding genes or transcription factors, but rather involve regulatory changes that are flexible enough to allow the large number of origins and reversals of the two types of pollination syndromes.

Candidate genes

The morphological changes associated with shifts in pollination syndromes in Gesneriaceae family include modifications in shape, size, and color of the corolla, nectar concentration, size of the reproductive organs, among others (Perret *et al.*, 2001; Schulte *et al.*, 2015; Serrano-Serrano *et al.*, 2015). For the investigated species, bee-pollinated flowers display larger corolla opening with a landing platform, inserted reproductive organs, and usually whitish to light purple colors (Fig. 3 and 4). In contrast, hummingbird-pollinated flowers are tubular with closed corolla mouth, exerted reproductive organs and bright reddish or pink corollas. Few studies have investigated the genetic control of pollination syndrome morphologies in the Gesneriaceae family. Alexandre *et al.* (2015) surveyed second-generation population of hybrids between two *Rhytidophyllum* species and identified a single QTLs associated with color, three QTLs associated with corolla shape differences and pollination syndromes, but few candidate genes were drawn from this evaluation. Here, we discussed the identified genetic responses in these *Gesneriaceae* species (Supp. Mat. S1 on line), and compared with the previously recognized genes in model systems (Galliot *et al.*, 2006). We will discuss three main morphological changes: corolla and organ size, corolla shape or symmetry, and corolla color.

Corolla and organ size

Studies on the genetic control of flower size indicated that changes are achieved through cell proliferation, vacuolization, cell wall loosening and expansion (Krizek & Anderson, 2013).We found four DEG involved in the regulation of cell division, the first belonging to the growth-regulating factors family (GRF-9, see Supp. Mat. S1) and three others related to the ABCE model (AP1 class A, and three AGAMOUS-like AGL9, AGL15, and AGL61). Only AGL15 showed up-regulation at the floral stage in

bee-pollinated species, while AP1 and AGL9 have patterns associated with vegetative tissues. AGL61 showed signatures of positive selection in bee-pollinated species, which has been hypothesized for other AGAMOUS-like genes during the evolution of floral variation in Zingiberales flowers (Almeida *et al.*, 2015).

Regarding cell expansion, we found the up-regulation of an expansin-15 gene in adult flowers of hummingbird-pollinated species (Fig. 8). Similar genes seem to mediate cell expansion and the increase in size of petal limbs in *Petunia* plants (Zenoni *et al.*, 2011). Cell expansion/elongation is also responsible for exserted stigmas in cross-pollinated cultivated tomatoes (Chen *et al.*, 2007), which is one of the key traits in the investigated hummingbird-pollinated species. Finally, genes associated with cell wall structure and loosening were found under purifying selection in hummingbird-pollinated species (SBT1.7, EP1, and VP52A). All these proteins could represent relevant genetic associations with the investigated floral morphologies.

Corolla shape and symmetry

All studied species have zygomorphic corollas, involving top and bottom differential development, and dorsoventrality. The genetic control of corolla shape and symmetry has been studied in *Antirrhinum* species and other groups such as Brassicales, Dipsacales, Malpighiales and Zingiberales, with the identification of CYC-like genes, such as CYC, DICH, DIV, and other TCP and B-class genes (Corley *et al.*, 2005; Preston *et al.*, 2011; Preston & Hileman, 2012; Glover, 2014). We identified no signal of DE from this group of genes, and only purifying selection was found in bee-pollinated species for the TCP15 gene (Supp. Mat. S1). This gene has two possible roles, first as repressor of style and stigma development (Lucero *et al.*, 2015), and as a negative regulator of anthocyanin accumulation in leaves of *Arabidopsis* (Viola et al., 2016). These roles point to the associated traits of bee-pollinated flowers, and the purifying selection could indicate constraints in this regulatory gene, in order to keep these functions. A transcriptomic study in *Orchis* also found that purifying selection was acting on multiple TCP transcripts (De Paolo *et al.*, 2015), but further investigation of this gene and the whole TCP family of TF in Gesneriaceae, is required to elucidate its functional role and sequence evolution. Finally, there is evidence that petal shape and curvature, producing traits such as landing platforms for bee-pollination, can be influenced by genes affecting cell differentiation, microtubule organization, and growth of lobes (Crawford *et al.*, 2004; Komaki & Sugimoto, 2012). We identified a variety of genes showing patterns of positive selection in bee-pollinated species that could play a role in these functions.

Corolla color

One of the most striking differences between flowers is dictated by color. Corolla colors are given by three types of plant pigments: flavonoids/anthocyanins (major pigments for orange to blue colors), betalains (yellow to red found in Caryophyllales), and carotenoids (yellow to red, see Tanaka *et al.*, 2008). Flavonoids are water soluble pigments, accumulated in cell vacuoles, with the anthocyanins being one of the most abundant flavonoid pigment in many plant tissues. The anthocyanin biosynthetic pathway is well understood (Tanaka *et al.*, 2008), and the genes involved in the synthesis of this pigment have been investigated in multiple plant groups (Bradshaw & Schemske, 2003; Smith *et al.*, 2013; Wessinger & Rausher, 2014). Few studies have explored the floral pigments in the Gesneriaceae family, however preliminary evidence suggest that *Sinningia* species may contain two anthocyanins and three flavone glycosides (Winefield *et al.*, 2005). Unpublished analysis in two of the sampled species indicate that cyanidin, delphinidin, and apigeninidin are present in *P. tenuiflora* and *S. magnifica* species (Master project conducted by Céline Caseys, at CJB).

We used previous studies to hypothesize the regulatory and structural regions that may act coordinately

in the genetic control of floral colors. The genes contributing to regulation of the pathway belong mainly to three groups of TFs: the R2R3-MYB, bHLH, and WD40-repeat types (Sheehan *et al.*, 2012; Sobel & Streisfeld, 2013). Here, we identified nine MYB TFs, seven of them recognized as R2R3, and four from the bHLH type (Table 4). Three MYB TFs are up-regulated in bee-pollinated species, with a reported potential role in volatile production, responses to UV stress, and trichome development (Myb86, Myb1, and Myb34 [see fig. 5.7], in Albert *et al.*, 2014; Zhao *et al.*, 2015). In contrast, the Myb305 gene was identified with an increase in expression in hummingbird-pollinated flowers at early stages. This gene has been previously suggested to activate the transcription of anthocyanin genes (Moyano *et al.*, 1996).

Several structural genes, for example F3H, F3PH, DFR, and ANS, have shown evidence of differential expression and sequence degeneration during color transitions (Smith *et al.*, 2013; Wessinger & Rausher, 2014). Surprisingly, very few structural genes in the anthocyanin pathway were identified in our study. First, the flavonoid 3'-monooxygenase (F3PH) was found up-regulated at stages B2 and flower (Fig. 8), and under signatures of purifying selection in bee-pollinated species. This gene is responsible for converting Dihydrokaempferol into Dihydroquercetin, and potentially deviating the metabolic pathway towards non-red anthocyanins (see Fig. 1 in Wessinger & Rausher, 2014), and other flavanones (e.g. Eriodictyol, see Fig. 1 in Sharma *et al.*, 2012). This deviation seems quite plausible as the flavonoids and anthocyanins share part of their biosynthesis pathway and substrate competition can be common (Yuan *et al.*, 2016). The second gene associated with the anthocyanin pathway is the Detoxification 41 protein (DTX41, or Transparent Testa TT12), which was found up-regulated in hummingbird-pollinated species in all developmental stages (Fig. 8). Previous reports suggested that this protein acts as a vacuolar antiporter of flavonoids in proanthocyanidin-accumulating cells in *A. thaliana* seed coat (Debeaujon *et al.*, 2001). Proanthocyanidins are colorless polyphenols commonly found in seed coats, and producing brown pigments after oxidation (Park *et al.*, 2007). These pigments and anthocyanins can be spatially co-localized having recruited genetic mechanisms (Abeynayake *et al.*, 2012). It can thus be hypothesized that the TT12 gene is used in a similar way in the Gesneriaceae to allocate or accumulate colored pigments in the hummingbird-pollinated species. Genes such as DFR and ANS were not identified by our analyses, but it is worthwhile to explore their specific expression and sequence evolution, in light of the patterns previously found (Whittall *et al.*, 2006; Smith *et al.*, 2013).

Conclusions

This chapter brings the first evaluation of the concerted genetic control underlying the parallel evolution of pollination syndromes in Gesneriaceae flowers. It provides novel information on the role of gene expression and coding sequence changes in floral morphology, and moreover linking with chapter 2, a broader view on the mechanisms for speciation in the family (Butlin *et al.*, 2012). We used the transcriptomic information to analyze the changes in gene expression and selection signatures during the transitions between bee and hummingbird pollination in three pairs of closely related species. We found a larger amount of genes showing common sequence patterns (i.e. selection signatures) than shared expression patterns. The amount of concerted DEG between pollination types increased with developmental time, suggesting that later stages of flower development share more genetic control than early expression profiles. Our investigation of the genetic control of hummingbird-pollinated flowers showed few shared DEG and many genes that have signature of diversifying selection. These results suggest that the evolution of the hummingbird-pollinated phenotype in the Gesneriaceae family is produced by highly variable genetic mechanisms in contrast to well-known study systems. The existence of multiple possible genetic solutions to produce the typical hummingbird-pollinated flower seem to contribute to the lability of the evolution of this phenotype in the Gesneriaceae, recalling the idea that few shared elements can underlay the convergent evolution of a trait, though they tend to overlap at the

pathway and biological function level (Berens *et al.*, 2015). Overall these trends reflect the diverse mechanisms for replicated adaptive responses (Elmer & Meyer, 2011), especially in systems with a large amount of evolutionary transitions, as it has been found for color production in Solanaceae and Antirrhineae flowers (Ellis & Field, 2016; Ng & Smith, 2016).

Our examination followed the promising offer of high throughput genomic studies to understand the natural variation in floral traits (Hermann & Kuhlemeier, 2011). However, under the macroevolutionary scenarios investigated, it is relevant to consider that the genetic mechanisms leading to parallel ecological adaptations (e.g. plant-pollinator interactions) may be unlikely to reveal shared patterns, due to differences in population demographic histories, species divergence times, and genetic architectures (Laurent *et al.*, 2016). Our model clade approach is very strict for picking up genes with a generalized effect in all the pairs of species investigated. The non-identification of genes previously associated with floral phenotypes can be the result of difficulties in finding orthologous genes via *de novo* comparative transcriptomics. Our study, also illustrate the fact that phenotypic responses are non-exclusively associated with expression changes in the expected metabolic pathways (as shown for corals in Barshis *et al.*, 2013), or that mRNA products and protein levels may not be completely correlated (Wu *et al.*, 2014). Further, post-transcriptional regulation could contribute to the discordances between gene expression, protein levels and phenotypic effects (Vogel & Marcotte, 2012; Velez-Bermudez & Schmidt, 2014).

Finally, many of the identified candidate genes may interact in multiple metabolic pathways, especially for regulatory regions in floral coloration and trichome development, among others. This result suggest that pleiotropic effects could underlay the phenotypic responses, as shown in other systems (Hermann *et al.*, 2013; Sheehan *et al.*, 2016). Our findings provide a first step towards a more extensive investigation of the phenotypic variation, trait correlations and genetic constraints in the Gesneriaceae family.

To summarize our results about the comparative transcriptomic between the six species, and potentially the genetic basis of the convergent pollination syndromes, indicated that: i) there are more concerted responses at the sequence than expression level. ii) Bee-pollinated species showed more concerted up-regulated genes and larger signals of purifying selection, suggesting conserved genetic pathways and constraints during their evolution. iii) In contrast, the number of concerted responses in gene expression is lower, and the signatures of positive selection larger for hummingbird-pollinated species, pointing to alternative ways to produce hummingbird-like flowers, with divergent selection acting during the evolution of this phenotype.

Consolidating our results about the candidate genes, we identified multiple genes that potentially play a role in the genetic control of the floral morphologies investigated. Many of them correlated with previously identified biochemical pathways, and are likely to be the ground of future macro-evolutionary studies. Other genes suggested alternative routes for controlling the observed changes between species, encouraging for more detailed experimental analyses. The genes that were expected to be identified but not revealed in our analyses may indicate the absence of parallel genetic responses in the system, and may reflect the diversity of the genetic mechanisms between explored transitions events.

Table 4. Transcription factors identified in the expression and sequence analyses as relevant for pollination types.

Gene name	Signal	Observations
Myb44	Purifying selection in B-pollinated	R2R3-type genes, seems associated with multicellular trichome development in *Cucumis sativus* (Zhao *et al.*, 2015)
Myb08	Positive selection in B-pollinated	R2R3-MYB genes (from Uniprot)
Myb308	Positive selection in B-pollinated. Showed expression in all plant organs, increasing as flowers develop	(Jackson *et al.*, 1991)
Myb05/305	Up-regulated in H-pollinated (B1)	Myb-related protein 305 is of potential insight into the flavonoid regulation, it seems to activate transcription of PAL, CHI and F3H (Moyano *et al.* 1996) biosynthesis. Expression increased as flowers develop. Trichome development (Zhao *et al.* 2015)
Myb39	Positive selection in B-pollinated	R2R3-MYB genes
Myb-APL	Purifying selection in B-pollinated	
Myb86	Up-regulated in B-pollinated (B2, FL, VG)	R2R3-type genes, also called myb4, fine-tunes the floral volatile, signature of *Petunia x hybrida* through PhC4H (Colquhoun *et al.*, 2010) TF in trichome development (Zhao *et al.* 2015)
Myb1	Up-regulated in B-pollinated (VG)	R2R3-type genes (from Uniprot)
Myb34	Up-regulated in B-pollinated (all stages)	Also called ATR1 activates tryptophan gene expression in *Arabidopsis thaliana* (from Uniprot)
bHLH75	Up-regulated in H-pollinated (B1,B2,VG)	
bHLH130	Purifying selection in B-pollinated	Associated with multicellular trichome development in *Cucumis sativus* (Zhao *et al.* 2015)
bHLH93	Positive selection in H-pollinated	
bHLH140	Positive selection in H-pollinated	
ERF025	Up-regulated in B-pollinated (early stages)	Associated with multicellular trichome development in *Cucumis sativus* (Zhao *et al.* 2015)
ERF110	Up-regulated in B-pollinated (B2 and flower stages)	
ERF003	Up-regulated in B-pollinated (all stages)	
ERF114	Up-regulated in H-pollinated (B1)	
ERF001	Up-regulated in B-pollinated (B1)	

Supplementary material on line S1:

https://www.dropbox.com/s/ax0uz52li5p13ry/Summary_DEG_selection.xlsx?dl=0

References

Abeynayake SW, Panter S, Chapman R, Webster T, Rochfort S, Mouradov A, Spangenberg G. 2012. Biosynthesis of proanthocyanidins in white clover flowers: cross talk within the flavonoid pathway. *Plant Physiol* **158**(2): 666-678.

Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. *Plant Cell* **26**(3): 962-980.

Alcantara S, Lohmann LG. 2010. Evolution of floral morphology and pollination system in Bignonieae (Bignoniaceae). *Am J Bot* **97**(5): 782-796.

Alexandre H, Vrignaud J, Mangin B, Joly S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus Rhytidophyllum (Gesneriaceae). *PeerJ* **3**: e1028.

Almeida AM, Yockteng R, Otoni WC, Specht CD. 2015. Positive selection on the K domain of the AGAMOUS protein in the Zingiberales suggests a mechanism for the evolution of androecial morphology. *Evodevo* **6**: 7.

Barshis DJ, Ladner JT, Oliver TA, Seneca FO, Traylor-Knowles N, Palumbi SR. 2013. Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences* **110**(4): 1387-1392.

Bartlett ME, Specht CD. 2010. Evidence for the involvement of Globosa-like gene duplications and expression divergence in the evolution of floral morphology in the Zingiberales. *New Phytol* **187**(2): 521-541.

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J. 2006. Positive selection driving diversification in plant secondary metabolism. *Proc Natl Acad Sci U S A* **103**(24): 9118-9123.

Berens AJ, Hunt JH, Toth AL. 2015. Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Mol Biol Evol* **32**(3): 690-703.

Bradshaw H, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* **426**(6963): 176-178.

Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M. 2011. The evolution of gene expression levels in mammalian organs. *Nature* **478**(7369): 343-348.

Brown JH, Kodric-Brown A. 1979. Convergence, competition, and mimicry in a temperate community of hummingbird-pollinated flowers. *Ecology* **60**(5): 1022-1035.

Butlin R, Debelle A, Kerth C, Snook RR, Beukeboom LW, Castillo Cajas RF, Diao W, Maan ME, Paolucci S, Weissing FJ, et al. 2012. What do we need to know about speciation? *Trends Ecol Evol* **27**(1): 27-39.

Castellanos MC, Wilson P, Thomson JD. 2004. 'Anti-bee' and 'pro-bird' changes during the evolution of hummingbird pollination in Penstemon flowers. *J Evol Biol* **17**(4): 876-885.

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE. 2016. Evolving Ideas on the Origin and Evolution of Flowers: New Perspectives in the Genomic Era. *Genetics* **202**(4): 1255-1265.

Chen K-Y, Cong B, Wing R, Vrebalov J, Tanksley SD. 2007. Changes in regulation of a transcription factor lead to autogamy in cultivated tomatoes. *Science* **318**(5850): 643-645.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome biology* **14**(6): 1.

Colquhoun TA, Kim JY, Wedde AE, Levin LA, Schmitt KC, Schuurink RC, Clark DG. 2010. PhMYB4 fine-tunes the floral volatile signature of Petunia× hybrida through PhC4H. *Journal of experimental botany*: erq342.

Corley SB, Carpenter R, Copsey L, Coen E. 2005. Floral asymmetry involves an interplay between TCP and MYB transcription factors in Antirrhinum. *Proc Natl Acad Sci U S A* **102**(14): 5068-5073.

Crawford BC, Nath U, Carpenter R, Coen ES. 2004. CINCINNATA controls both cell differentiation and growth in petal lobes and leaves of Antirrhinum. *Plant Physiology* **135**(1): 244-253.

Cronk Q, Ojeda I. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *J Exp Bot* **59**(4): 715-727.

De Paolo S, Gaudio L, Aceto S. 2015. Analysis of the TCP genes expressed in the inflorescence of the orchid Orchis italica. *Scientific reports* **5**.

Debeaujon I, Peeters AJ, Léon-Kloosterziel KM, Koornneef M. 2001. The TRANSPARENT TESTA12 gene of Arabidopsis encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *The Plant Cell* **13**(4): 853-871.

DeBiasse MB, Kelly MW. 2016. Plastic and Evolved Responses to Global Change: What Can We Learn from Comparative Transcriptomics? *J Hered* **107**(1): 71-81.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in Ipomoea. *Evolution* **64**(7): 2044-2054.

Ellis TJ, Field DL. 2016. Repeated gains in yellow and anthocyanin pigmentation in flower colour transitions in the Antirrhineae. *Ann Bot* **117**(7): 1133-1140.

Elmer KR, Meyer A. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol Evol* **26**(6): 298-306.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination Syndromes and Floral Specialization. *Annual Review of Ecology, Evolution, and Systematics* **35**(1): 375-403.

Foote AD, Liu Y, Thomas GW, Vinar T, Alfoldi J, Deng J, Dugan S, van Elk CE, Hunter ME, Joshi V, et al. 2015. Convergent evolution of the genomes of marine mammals. *Nat Genet* **47**(3): 272-275.

Galliot C, Stuurman J, Kuhlemeier C. 2006. The genetic dissection of floral pollination syndromes. *Curr Opin Plant Biol* **9**(1): 78-82.

Glover B. 2014. *Understanding Flowers and Flowering Second Edition*: OUP Oxford.

Guindon S, Delsuc F, Dufayard J-F, Gascuel O. 2009. Estimating maximum likelihood phylogenies with PhyML. *Bioinformatics for DNA sequence analysis*: 113-137.

Harder LD, Johnson SD. 2009. Darwin's beautiful contrivances: evolutionary and functional evidence for floral adaptation. *New Phytol* **183**(3): 530-545.

Harrison PW, Wright AE, Mank JE. 2012. The evolution of gene expression and the transcriptome-phenotype relationship. *Semin Cell Dev Biol* **23**(2): 222-229.

Hermann K, Klahre U, Moser M, Sheehan H, Mandel T, Kuhlemeier C. 2013. Tight genetic linkage of prezygotic barrier loci creates a multifunctional speciation island in Petunia. *Curr Biol* **23**(10): 873-877.

Hermann K, Kuhlemeier C. 2011. The genetic architecture of natural variation in flower morphology. *Curr Opin Plant Biol* **14**(1): 60-65.

Hileman LC. 2014. Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philos Trans R Soc Lond B Biol Sci* **369**(1648).

Hodgins KA, Yeaman S, Nurkowski KA, Rieseberg LH, Aitken SN. 2016. Expression divergence is correlated with sequence evolution but not positive selection in conifers. *Molecular Biology and Evolution*: msw032.

Hunt BG, Ometto L, Keller L, Goodisman MA. 2012. Evolution at two levels in fire ants: the relationship between patterns of gene expression and protein sequence evolution. *Molecular Biology and Evolution*: mss234.

Jackson D, Culianez-Macia F, Prescott AG, Roberts K, Martin C. 1991. Expression patterns of myb genes from Antirrhinum flowers. *The Plant Cell* **3**(2): 115-125.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**(4): 772-780.

Kay KM, Schemske DW. 2003. Pollinator assemblages and visitation rates for 11 species of Neotropical Costus (Costaceae). *Biotropica* **35**(2): 198-207.

Kevan P, Baker H. 1983. Insects as flower visitors and pollinators. *Annual review of entomology* **28**(1): 407-453.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Paabo S. 2004. A neutral model of transcriptome evolution. *PLoS Biol* **2**(5): E132.

Komaki S, Sugimoto K. 2012. Control of the plant cell cycle by developmental and environmental cues. *Plant and Cell Physiology* **53**(6): 953-964.

Krizek BA, Anderson JT. 2013. Control of flower size. *J Exp Bot* **64**(6): 1427-1437.

Laurent S, Pfeifer SP, Settles ML, Hunter SS, Hardwick KM, Ormond L, Sousa VC, Jensen JD, Rosenblum EB. 2016. The population genomics of rapid adaptation: disentangling signatures of selection and demography in white sands lizards. *Mol Ecol* **25**(1): 306-323.

Law CW, Chen Y, Shi W, Smyth GK. 2014. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome biology* **15**(2): 1.

Lucero LE, Uberti-Manassero NG, Arce AL, Colombatti F, Alemano SG, Gonzalez DH. 2015. TCP15 modulates cytokinin and auxin responses during gynoecium development in Arabidopsis. *Plant J* **84**(2): 267-282.

Marten-Rodriguez S, Fenster CB, Agnarsson I, Skog LE, Zimmer EA. 2010. Evolutionary breakdown of pollination specialization in a Caribbean plant radiation. *New Phytol* **188**(2): 403-417.

Martin A, Orgogozo V. 2013. The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**(5): 1235-1250.

McGhee GR. 2011. *Convergent evolution: limited forms most beautiful*: MIT Press.

Michez D, Vanderplanck M, Engel MS, Patiny S 2012. Fossil bees and their plant associates. In: Patiny S ed. *Evolution of plant-pollinator relationships*, 103-164.

Moyano E, Martínez-Garcia JF, Martin C. 1996. Apparent redundancy in myb gene function provides gearing for the control of flavonoid biosynthesis in Antirrhinum flowers. *The Plant Cell* **8**(9): 1519-1532.

Ng J, Smith SD. 2016. Widespread flower color convergence in Solanaceae via alternate biochemical pathways. *New Phytol* **209**(1): 407-417.

Ord TJ, Summers TC. 2015. Repeated evolution and the impact of evolutionary history on adaptation. *BMC Evol Biol* **15**: 137.

Pankey MS, Minin VN, Imholte GC, Suchard MA, Oakley TH. 2014. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proceedings of the National Academy of Sciences* **111**(44): E4736-E4742.

Park KI, Ishikawa N, Morita Y, Choi JD, Hoshino A, Iida S. 2007. A bHLH regulatory gene in the common morning glory, Ipomoea purpurea, controls anthocyanin biosynthesis in flowers, proanthocyanidin and phytomelanin pigmentation in seeds, and seed trichome formation. *Plant J* **49**(4): 641-654.

Pavey SA, Collin H, Nosil P, Rogers SM. 2010. The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences* **1206**(1): 110-129.

Perez-Rodriguez M, Jaffe FW, Butelli E, Glover BJ, Martin C. 2005. Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of Antirrhinum majus flowers. *Development* **132**(2): 359-370.

Perret M, Chautems A, Spichiger R, Barraclough TG, Savolainen V. 2007. The geographical pattern of speciation and floral diversification in the neotropics: the tribe Sinningieae (Gesneriaceae) as a case study. *Evolution* **61**(7): 1641-1660.

Perret M, Chautems A, Spichiger R, Peixoto M, Savolainen V. 2001. Nectar sugar composition in relation to pollination syndromes in Sinningieae (Gesneriaceae). *Annals of Botany* **87**(2): 267-273.

Pires ND, Dolan L. 2012. Morphological evolution in land plants: new designs with old genes. *Philos Trans R Soc Lond B Biol Sci* **367**(1588): 508-518.

Preston JC, Hileman LC. 2012. Parallel evolution of TCP and B-class genes in Commelinaceae flower bilateral symmetry. *Evodevo* **3**: 6.

Preston JC, Hileman LC, Cubas P. 2011. Reduce, reuse, and recycle: developmental evolution of trait diversification. *Am J Bot* **98**(3): 397-403.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**(7): e47.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1): 139-140.

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**(3): R25.

Rohlfs RV, Harrigan P, Nielsen R. 2014. Modeling gene expression evolution with an extended Ornstein-Uhlenbeck process accounting for within-species variation. *Mol Biol Evol* **31**(1): 201-211.

Rosas-Guerrero V, Aguilar R, Marten-Rodriguez S, Ashworth L, Lopezaraiza-Mikel M, Bastida JM, Quesada M. 2014. A quantitative review of pollination syndromes: do floral traits predict effective pollinators? *Ecol Lett* **17**(3): 388-400.

Rosenblum EB, Parent CE, Brandt EE. 2014. The Molecular Basis of Phenotypic Convergence. *Annual Review of Ecology, Evolution, and Systematics* **45**(1): 203-226.
SanMartin-Gajardo I, Sazima M. 2004. Non-Euglossine bees also function as pollinators of Sinningia species (Gesneriaceae) in southeastern brazil. *Plant Biol (Stuttg)* **6**(4): 506-512.
SanMartin-Gajardo I, Sazima M. 2005. Chiropterophily in Sinningieae (Gesneriaceae): Sinningia brasiliensis and Paliavana prasinata are bat-pollinated, but P. sericiflora is not. Not yet? *Ann Bot* **95**(7): 1097-1103.
Schemske DW. 1981. Floral convergence and pollinator sharing in two bee-pollinated tropical herbs. *Ecology* **62**(4): 946-954.
Schulte LJ, Clark JL, Novak SJ, Jeffries SK, Smith JF. 2015. Speciation within Columnea section Angustiflora (Gesneriaceae): islands, pollinators and climate. *Mol Phylogenet Evol* **84**: 125-144.
Serrano-Serrano ML, Perret M, Guignard M, Chautems A, Silvestro D, Salamin N. 2015. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evol Biol* **15**: 247.
Sharma M, Chai C, Morohashi K, Grotewold E, Snook ME, Chopra S. 2012. Expression of flavonoid 3'-hydroxylase is controlled by P1, the regulator of 3-deoxyflavonoid biosynthesis in maize. *BMC plant biology* **12**(1): 1.
Sheehan H, Hermann K, Kuhlemeier C. 2012. Color and scent: how single genes influence pollinator attraction. *Cold Spring Harb Symp Quant Biol* **77**: 117-133.
Sheehan H, Moser M, Klahre U, Esfeld K, Dell'Olivo A, Mandel T, Metzger S, Vandenbussche M, Freitas L, Kuhlemeier C. 2016. MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. *Nat Genet* **48**(2): 159-166.
Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a flower color transition. *Mol Biol Evol* **30**(3): 602-612.
Sobel JM, Streisfeld MA. 2013. Flower color as a model system for studies of plant evo-devo. *Front Plant Sci* **4**: 321.
Soneson C, Delorenzi M. 2013. A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* **14**(1): 1.
Steiner CC, Römpler H, Boettger LM, Schöneberg T, Hoekstra HE. 2009. The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Molecular Biology and Evolution* **26**(1): 35-45.
Strauss SY, Whittall JB. 2006. Non-pollinator agents of selection on floral traits. *Ecology and evolution of flowers*: 120-138.
Stuurman J, Hoballah ME, Broger L, Moore J, Basten C, Kuhlemeier C. 2004. Dissection of floral pollination syndromes in Petunia. *Genetics* **168**(3): 1585-1599.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic acids research* **34**(suppl 2): W609-W612.
Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J* **54**(4): 733-749.
Thomson James D, Wilson P. 2008. Explaining Evolutionary Shifts between Bee and Hummingbird Pollination: Convergence, Divergence, and Directionality. *International Journal of Plant Sciences* **169**(1): 23-38.
Tripp EA, Manos PS. 2008. Is floral specialization an evolutionary dead-end? Pollination system transitions in Ruellia (Acanthaceae). *Evolution* **62**(7): 1712-1737.
Velez-Bermudez IC, Schmidt W. 2014. The conundrum of discordant protein and mRNA expression. Are plants special? *Front Plant Sci* **5**: 619.
Viola IL, Camoirano A, Gonzalez DH. 2016. Redox-Dependent Modulation of Anthocyanin Biosynthesis by the TCP Transcription Factor TCP15 during Exposure to High Light Intensity Conditions in Arabidopsis. *Plant Physiol* **170**(1): 74-85.
Vogel C, Marcotte EM. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* **13**(4): 227-232.
Warnefors M, Kaessmann H. 2013. Evolution of the correlation between expression divergence and protein divergence in mammals. *Genome Biol Evol* **5**(7): 1324-1335.
Weadick CJ, Chang BS. 2012. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Mol Biol Evol* **29**(5): 1297-1300.
Weiss MR. 1995. Floral color change: a widespread functional convergence. *American Journal of Botany*: 167-185.

Wessinger CA, Hileman LC, Rausher MD. 2014. Identification of major quantitative trait loci underlying floral pollination syndrome divergence in Penstemon. *Philos Trans R Soc Lond B Biol Sci* **369**(1648).

Wessinger CA, Rausher MD. 2014. Predictability and irreversibility of genetic changes associated with flower color evolution in Penstemon barbatus. *Evolution* **68**(4): 1058-1070.

Whittall JB, Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* **447**(7145): 706-709.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in Aquilegia. *Mol Ecol* **15**(14): 4645-4657.

Wilson P, Wolfe AD, Armbruster WS, Thomson JD. 2007. Constrained lability in floral evolution: counting convergent origins of hummingbird pollination in Penstemon and Keckiella. *New Phytol* **176**(4): 883-890.

Winefield CS, Lewis DH, Swinny EE, Zhang H, Arathoon HS, Fischer TC, Halbwirth H, Stich K, Gosch C, Forkmann G. 2005. Investigation of the biosynthesis of 3-deoxyanthocyanins in Sinningia cardinalis. *Physiologia Plantarum* **124**(4): 419-430.

Woodard SH, Fischman BJ, Venkat A, Hudson ME, Varala K, Cameron SA, Clark AG, Robinson GE. 2011. Genes involved in convergent evolution of eusociality in bees. *Proceedings of the National Academy of Sciences* **108**(18): 7472-7477.

Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics* **8**(3): 206-216.

Wu Y, Williams EG, Dubuis S, Mottis A, Jovaisaite V, Houten SM, Argmann CA, Faridi P, Wolski W, Kutalik Z. 2014. Multilayered genetic and omics dissection of mitochondrial activity in a mouse reference population. *Cell* **158**(6): 1415-1430.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**(8): 1586-1591.

Yoo MJ, Chanderbali AS, Altman NS, Soltis PS, Soltis DE. 2010. Evolutionary trends in the floral transcriptome: insights from one of the basalmost angiosperms, the water lily Nuphar advena (Nymphaeaceae). *Plant J* **64**(4): 687-698.

Yuan Y-W, Rebocho AB, Sagawa JM, Stanley LE, Bradshaw HD. 2016. Competition between anthocyanin and flavonol biosynthesis produces spatial pattern variation of floral pigments between Mimulus species. *Proceedings of the National Academy of Sciences* **113**(9): 2448-2453.

Zenoni S, Fasoli M, Tornielli GB, Dal Santo S, Sanson A, de Groot P, Sordo S, Citterio S, Monti F, Pezzotti M. 2011. Overexpression of PhEXPA1 increases cell size, modifies cell wall polymer composition and affects the timing of axillary meristem development in Petunia hybrida. *New Phytol* **191**(3): 662-677.

Zhang L, Yan H-F, Wu W, Yu H, Ge X-J. 2013. Comparative transcriptome analysis and marker development of two closely related Primrose species (Primula poissonii and Primula wilsonii). *BMC Genomics* **14**(1): 1.

Zhang W, Kramer EM, Davis CC. 2012. Similar genetic mechanisms underlie the parallel evolution of floral phenotypes. *PLoS One* **7**(4): e36033.

Zhao JL, Pan JS, Guan Y, Nie JT, Yang JJ, Qu ML, He HL, Cai R. 2015. Transcriptome analysis in Cucumis sativus identifies genes involved in multicellular trichome development. *Genomics* **105**(5-6): 296-303.

Zufall RA, Rausher M. 2003. The genetic basis of a flower color polymorphism in the common morning glory (Ipomoea purpurea). *Journal of Heredity* **94**(6): 442-448.

Zwiener I, Frisch B, Binder H. 2014. Transforming RNA-Seq data to improve the performance of prognostic gene signatures. *PLoS One* **9**(1): e85150.

Synthesis and perspectives

The advances in understanding the species generation process have shifted from characterizing uniquely the geographical context of related species, towards an integrative research that connects the macro-evolutionary patterns of biodiversity, geographical and ecological backgrounds, with the building-up of reproductive isolation and the associated genomic changes (Wiens, 2011; Butlin *et al.*, 2012). Numerous reviews have highlighted the need for an evaluation of the interacting agents in plant speciation (Antonelli & Sanmartín, 2011; Vamosi & Vamosi, 2011), and the nature and effect of the genetic elements involved in the process (Lexer & Widmer, 2008). In my thesis I integrated several of these aspects to better understand biodiversity patterns in the Neotropical lineage of the Gesneriaceae family, especially in the light of plant-pollinator interactions. Gesneriaceae represent an excellent study system because they are species-rich, and very morphologically and ecologically diverse. Thus, I used this family to bridge ecological and evolutionary theories, for a significant understanding of the effects of plant-pollinator interactions in Neotropical biodiversity. I combined macroevolutionary and genomic methods with multiple evolutionary timescales. Each of these scales have advantages and limitations, and provide the ground for answering future questions on how biodiversity originates. Chapters 1 and 2 deal mostly with macroevolutionary analyses and phylogenetic methods, and chapters 3 and 4 focus on a transcriptomic evaluation of closely related species.

Understanding species diversification in the Gesneriaceae

In chapter 1 I have demonstrated that the evolution of morphology and climatic preferences, in 46 species within the genera *Codonanthe* and *Nematanthus*, is decoupled. My findings indicated that the evolution of floral shape is constrained by the plant-pollinator interaction. In contrast, floral size showed an increase in the rate of evolution in a hummingbird-pollinated clade. Climatic preferences are highly dependent on the separate biome distribution, with precipitation regime preferences evolving faster than any other climatic component (Serrano-Serrano *et al.*, 2015). This chapter also emphasized the importance of heterogeneous macroevolutionary models to investigate complex trait evolution scenarios and highlighted the advantage of using models of trait evolution with an explicit selective regime. Finally, within this analytical framework I developed new tools for the visualization of trait evolution to facilitate the interpretation of complex evolutionary dynamics. These dynamics go beyond the traditional Brownian motion model that assumes neutral evolution of a trait, to the Ornstein-Uhlenbeck models that incorporate selection and adaptive processes. Altogether, my findings in chapter 1 indicated that pollinators played a key role in the evolution of two genera of Gesneriaceae, prompting the question of whether consistent patterns can be found across the entire family.

In chapter 2 I explored the dynamics of pollination interactions at broader taxonomic level (the entire Gesnerioideae subfamily) and deeper evolutionary time scale (several tens of Myr): Do specific pollinators lead to increased diversification rates and how often did pollinator transitions occur in the family? To address these questions, I reconstructed a large phylogeny and the pollination syndrome for about 600 species in the subfamily. Using cutting-edge phylogenetic comparative methods allowed me to perform a test to detect shifts in diversification rates due to shifts in pollination syndromes, or due to specialization into a single pollination type. I primarily showed that the specialization in hummingbird-pollination has triggered an increase in diversification across multiple lineages in the subfamily, but also that plant-pollination interactions are highly labile in the group, with frequent transitions between the two main types of pollinators (Serrano-Serrano et al. *in review*). Overall, chapter 2 pointed to an active role of hummingbird-pollination in the speciation process of the family.

The influence of this pollination syndrome on speciation has been concomitantly suggested for other plant families such as Campanulaceae and Bromeliaceae (Givnish *et al.*, 2014; Lagomarsino *et al.*, 2016), but the generalization of hummingbirds as drivers of plant speciation is still elusive and requires an understanding of the underlying mechanisms (Schmidt-Lebuhn *et al.*, 2007). Our evidence from chapter 1 suggested that hummingbird pollination leads to an increase in the rates of floral size evolution, which could enhance the opportunities for reproductive isolation and speciation, as indeed the findings in chapter 2 indicate. Future work could assess whether different pollinator types show different rates of morphological diversification at the family-wide level, and within other plant families. Previous studies found that derived plant-pollinator interactions produce floral morphologies that largely deviate from the ancestral forms, by breaking existing allometric relationships (Strelin *et al.*, 2016), or increasing color diversity (Muchhala *et al.*, 2014). However, carrying out these evaluations will face the difficulties of obtaining a great quantity of precise and comparable morphological data across different plant groups. Such endeavor could however provide essential insights to answer whether the specialization in hummingbird pollination consistently enhance the divergence of flowers, while using more efficiently the same pollinator resources, and contributing to the observed patterns of neotropical plant diversification.

The investigation of the factors promoting plant speciation has additionally hypothesized that contingent traits may influence diversification (Vamosi & Vamosi, 2011; Donoghue & Sanderson, 2015). So far, I have looked at the exclusive role of plant-pollinator interactions in shaping the biodiversity of the family, and additional factors remain unexplored. In particular it would be interesting to investigate the potential role of the growth habit (terrestrial or epiphytic), and geographical distribution of subclades across different biomes in shaping the diversification of species. I tried to explore these contingent factors, but conventional biogeographic and trait evolution methods explore independently the evolutionary history of each trait, and their potential effect on diversification, making a poor hypothesis testing approach for a joint effect of factors. Moreover, the current methods are highly sensitive to few state transition events (pseudo-replication issue pointed by Maddison & FitzJohn, 2015), which could be easily found for some traits. An optimal methodology should evaluate multiple traits simultaneously, allowing a particular combination of those (i.e. colonization of a new area, with an in situ new trait state) to open a new diversification process, as this may constitute a way to escape the competition for resources or previous density-dependent dynamics (as discussed in Etienne & Haegeman, 2012).

Our data allowed the characterization of pollination syndromes in the family, with the inclusion of a large number of species, but misses the precise information on plant-pollination interactions. Promising avenues of research include questions such as how do plant-pollination interactions influence the community assemblies (Sargent & Ackerly, 2008)? Hummingbird communities in the Andes have structured local compositions, highly phylogenetically clustered, with differences between lowland and high elevation regions (Graham *et al.*, 2009), but so far little effort has been placed in the joint assessment of plant and pollinator community structures (Pellissier *et al.*, 2012). The evaluation of the spatial patterns of plant and pollinator species distribution, as well as their morphological variation, will help to understand the mechanisms enabling species to coexist, share pollinators, while reducing heterospecific pollen placements (Sargent & Ackerly, 2008). These directions should be applied to several biome-specific radiations in the Gesneriaceae family (Perret *et al.*, 2013), while testing whether they reflect multiple instances of in situ efficient exploitation of pollinator resources.

Understanding the genetic control of floral diversification in the Gesneriaceae

The broad geographic and taxonomic range used in the first two chapters has the advantage of testing large-scale patterns of species diversity. My results gave clues about the mechanisms increasing plant speciation by studying the morphological trait variation, the diversification rates, and the evolvability of pollination systems in Gesneriaceae (chapters 1 and 2). However, these phylogenetic patterns may fail to detect small scale mechanisms involved in reproductive isolation. This isolation, typically prezygotic when pollinators act as a driver of diversity, usually requires substantial changes in flower morphologies. Understanding the mechanisms for generating diversity in flower morphology calls for an investigation of which genes control those changes. To this aim, the accessibility of the "-omics" era into non-model species played an important role in the experimental design of chapters 3 and 4 (Alvarez *et al.*, 2015; Todd *et al.*, 2016). **In chapter 3**, I described the design and generation of the next-generation sequencing data for six related species in the Gesnerioideae subfamily.

In chapter 4, I examined the molecular mechanisms of floral morphogenesis to shed light on the convergent nature of the floral morphologies associated with specific pollination syndromes, and the specific genes controlling the floral changes during shifts in plant-pollinator interactions. I developed the first next-generation sequencing data for the Gesneriaceae family to evaluate the concerted changes in gene expression and sequence evolution during shifts in pollination type in the Gesneriaceae that are associated with recurrent morphologies. The genomic data produced in these non-model organisms allowed me to explore the contribution of different genetic elements during the transitions in pollination type in Gesneriaceae, and to hypothesize the role of specific genes in plant speciation (Rieseberg & Blackman, 2010).

I demonstrated that concerted genetic responses between the three cases of pollination transitions are scarce. Despite evidence that pollinators from a given functional group consistently favor the appearance of similar floral traits (e.g. convergent corolla shape and color), the genetic controls involved in the evolution of those traits in closely related species can take alternative routes to produce similar morphologies. The work done in my thesis further provides a large body of results to better understand the genes controlling traits such as corolla shape, size and color in the Gesneriaceae family. I found few instances of previously identified genes, associated with the mentioned traits in model systems, and a handful of candidate genes deserving further evaluation.

My analysis of the molecular mechanisms underlying floral morphologies in the subfamily adds new perspectives to the study of plant speciation. The generation of transcriptomic resources in six non-model species allowed the characterization of comparative gene expression, but without any previous genomic characterization, it presents several difficulties such as the correct orthology assessment, incompleteness of annotation, and lack of evidence for genomic structural variation. Given that molecular features such as structural reorganization, level of synteny, fluctuations in gene copy number, genome duplication events, and transposable elements, are known to contribute to reproductive isolation and speciation (Baack *et al.*, 2015), a genome sequencing project in at least a pair of species with different pollination types, will facilitate the study of such features. In particular, the exploration of gene duplication level is associated with a potential role in diversifying floral morphology, as a way to circumvent existing genetic correlations (Cooley *et al.*, 2011; Wessinger & Hileman, 2016), or allowing for relaxed selection while preserving the ancestral functions in a highly evolvable system.

As discussed in the previous section, multiple plant traits can influence the diversification process by allowing the species to better cope with new ecological conditions (Vamosi & Vamosi, 2011). Future research in these traits could be favored by the existing transcriptomic data, through the evaluation of multi-locus SNP (Single Nucleotide Polymorphism) or gene capture approaches at the population level along an environmental gradient (Turner *et al.*, 2010). Similar approaches can be performed between sets of species under contrasting conditions, such as epiphytic growth, root tuber formation, and CAM metabolism. These evaluations could elucidate the genomic architecture of such additional traits, their limited number of originations in the family, and their ecological trade-offs, to better understand their combined contribution to species diversification.

Demonstrating that floral diversification plays a role in speciation would be greatly helped by the identification of which morphological attributes are changing during the evolution of a new species, and the contribution of those changes to the establishment of new plant-pollinator interactions. New phenotypic variants may produce changes in pollination preferences, as previously shown for model species (Schemske & Bradshaw, 1999). Hybrid species are a quick source of intermediate phenotypes and new gene combinations (Givnish, 2010), and the use of artificial or natural hybrids could help to experimentally recognize the phenotypic and genetic variations produced during the evolution of new forms, as well as genetic constraints of this process. The Gesneriaceae species seem to hybridize in their natural range (M. Wolowski, personal communication, and based on my own observations during collaborative field work in 2015), and in greenhouse conditions (Chautems, 1988). These natural or artificial hybrids could bring the opportunity of developing quantitative trait locus and transgenic analyses, and conducting experimental evolution studies. However, such studies present important challenges on the propagation, maintenance and phenotypic evaluation (e.g. long flowering times) when working in non-model organisms. Even so, such studies have been recently pursued in the family with promising results (Alexandre *et al.*, 2015).

Few previous studies evaluating floral variation within hummingbird-pollinated species pointed to a rapid evolution of traits such as corolla pigmentation and size (Muchhala *et al.*, 2014; Serrano-Serrano *et al.*, 2015). These traits are, in some cases, known to be controlled by particular biochemical and functional pathways, and their investigation should lead to a deeper understanding of the genetic controls of flower morphologies and the possible routes for increased variation. Indeed, the symmetric and rapid rates of pollinator transitions in the family I inferred in chapter 2 offer a novel approach to assess whether similar or different mechanisms underlie this type of evolutionary event, in light of previously studied biological systems (e.g. *Ipomoea* and *Iochroma*). The sequence information generated from this work could help the exploration of gene expression at previously screened molecular pathways (e.g. pigment biosynthesis). The use of quantitative approaches like qPCR and functional assays (see examples in Des Marais & Rausher, 2010; Smith *et al.*, 2013) could represent a more precise characterization of these pathways, and profit of a larger sample of species from the *Sinningia* and *Nematanthus* genera, which are available at the Botanical Garden of Geneva.

Furthermore, quantitative approaches for gene expression could be expanded to a finer scale. It has been suggested that spatially discrete transcriptional profiles (i.e. organ-specific gene expression) can contribute to the evolution of derived angiosperm floral forms (Chanderbali *et al.*, 2010). Our results are based on samples composed by a mixture of floral tissues (see chapter 3), and detailed tissue- or spatial-specific gene expression (i.e. at particular locations in a flower tissue) could complement the understanding of flower diversity, especially for traits such as floral size where large morphological variability between species has been found (Serrano-Serrano *et al.* 2015). This approach should be guided

by a detailed morphological characterization, for instance of the patterns of cell proliferation and expansion, within a developmental framework in phenotypically different species.

Complementary to the evidence from patterns of gene expression, I showed that molecular evolution, specifically substitution rates and signatures of selection, has a strong effect on the repeated evolution of pollination syndrome (see chapter 4). It would be interesting to investigate whether these molecular patterns are also found during morphological evolution in a larger sample of species in the family. Similar questions have been addressed by the recent phylogenetic analyses of the large R2R3-MYB transcription factor family. Gates *et al.* (2016) have identified multiple taxon- and lineage-specific gain and losses of domains, and potential neo-functionalization during the floral evolution in the Solanaceae plant family. Thus, the integration of the identified candidate genes into a comparative framework, which would include a larger and representative sample of species, could help the identification of molecular mechanisms driving floral evolution into a larger evolutionary scale.

In conclusion the research projects I have developed in my thesis highlight the value of understanding the generation of floral divergences, the evolution of plant-pollinator interactions, and their combined effect on the increased speciation in the same plant group. I showed that the Gesneriaceae family is an excellent a system to unravel the genetic control of floral changes in a multi-species framework, and developed a comprehensive and flexible research framework to tackle evolutionary questions at different timescales. The combination of comparative and transcriptomic analyses have opened multiple ecological and evolutionary perspectives to elucidate the impact of hummingbird pollination in the building-up of biodiversity in the Neotropics.

References

Alexandre H, Vrignaud J, Mangin B, Joly S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus Rhytidophyllum (Gesneriaceae). *PeerJ* **3**: e1028.

Alvarez M, Schrey AW, Richards CL. 2015. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Molecular Ecology* **24**(4): 710-725.

Antonelli A, Sanmartín I. 2011. Why are there so many plant species in the Neotropics? *Taxon* **60**(2): 403-414.

Baack E, Melo MC, Rieseberg LH, Ortiz-Barrientos D. 2015. The origins of reproductive isolation in plants. *New Phytol* **207**(4): 968-984.

Butlin R, Debelle A, Kerth C, Snook RR, Beukeboom LW, Castillo Cajas RF, Diao W, Maan ME, Paolucci S, Weissing FJ, et al. 2012. What do we need to know about speciation? *Trends Ecol Evol* **27**(1): 27-39.

Chanderbali AS, Yoo M-J, Zahn LM, Brockington SF, Wall PK, Gitzendanner MA, Albert VA, Leebens-Mack J, Altman NS, Ma H. 2010. Conservation and canalization of gene expression during angiosperm diversification accompany the origin and evolution of the flower. *Proceedings of the National Academy of Sciences* **107**(52): 22570-22575.

Chautems A. 1988. Révision taxonomique et possibilités d'hybridations de Nematanthus Schrader (Gesneriaceae), genre endémique de la forêt côtière brésilienne. *Dissertationes botanicae* **112**.

Cooley AM, Modliszewski JL, Rommel ML, Willis JH. 2011. Gene duplication in Mimulus underlies parallel floral evolution via independent trans-regulatory changes. *Curr Biol* **21**(8): 700-704.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in Ipomoea. *Evolution* **64**(7): 2044-2054.

Donoghue MJ, Sanderson MJ. 2015. Confluence, synnovation, and depauperons in plant diversification. *New Phytol* **207**(2): 260-274.

Etienne RS, Haegeman B. 2012. A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *Am Nat* **180**(4): E75-89.

Gates DJ, Strickler SR, Mueller LA, Olson BJSC, Smith SD. 2016. Diversification of R2R3-MYB Transcription Factors in the Tomato Family Solanaceae. *Journal of Molecular Evolution*: 1-12.

Givnish TJ. 2010. Ecology of plant speciation. *Taxon* **59**(5): 1326-1366.

Givnish TJ, Barfuss MH, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JA, et al. 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Mol Phylogenet Evol* **71**: 55-78.

Graham CH, Parra JL, Rahbek C, McGuire JA. 2009. Phylogenetic structure in tropical hummingbird communities. *Proc Natl Acad Sci U S A* **106 Suppl 2**: 19673-19678.

Lagomarsino LP, Condamine FL, Antonelli A, Mulch A, Davis CC. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol* **210**(4): 1430-1442.

Lexer C, Widmer A. 2008. Review. The genic view of plant speciation: recent progress and emerging questions. *Philos Trans R Soc Lond B Biol Sci* **363**(1506): 3023-3036.

Maddison WP, FitzJohn RG. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* **64**(1): 127-136.

Muchhala N, Johnsen S, Smith SD. 2014. Competition for hummingbird pollination shapes flower color variation in Andean solanaceae. *Evolution* **68**(8): 2275-2286.

Pellissier L, Alvarez N, Guisan A. 2012. Pollinators as drivers of plant distribution and assemblage into communities. *Evolution of Plant-Pollinator Relationships. Systematics Association Special Volume Series*: 392-413.

Perret M, Chautems A, De Araujo AO, Salamin N. 2013. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. *Botanical Journal of the Linnean Society* **171**(1): 61-79.

Rieseberg LH, Blackman BK. 2010. Speciation genes in plants. *Ann Bot* **106**(3): 439-455.

Sargent RD, Ackerly DD. 2008. Plant-pollinator interactions and the assembly of plant communities. *Trends Ecol Evol* **23**(3): 123-130.

Schemske DW, Bradshaw H. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). *Proceedings of the National Academy of Sciences* **96**(21): 11910-11915.

Schmidt-Lebuhn AN, Kessler M, Hensen I. 2007. Hummingbirds as drivers of plant speciation? *Trends Plant Sci* **12**(8): 329-331.

Serrano-Serrano ML, Perret M, Guignard M, Chautems A, Silvestro D, Salamin N. 2015. Decoupled evolution of floral traits and climatic preferences in a clade of Neotropical Gesneriaceae. *BMC Evol Biol* **15**: 247.

Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a flower color transition. *Mol Biol Evol* **30**(3): 602-612.

Strelin MM, Benitez-Vieyra S, Fornoni J, Klingenberg CP, Cocucci AA. 2016. Exploring the ontogenetic scaling hypothesis during the diversification of pollination syndromes in Caiophora (Loasaceae, subfam. Loasoideae). *Ann Bot* **117**(5): 937-947.

Todd EV, Black MA, Gemmell NJ. 2016. The power and promise of RNA-seq in ecology and evolution. *Mol Ecol* **25**(6): 1224-1241.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV. 2010. Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. *Nature genetics* **42**(3): 260-263.

Vamosi JC, Vamosi SM. 2011. Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. *Am J Bot* **98**(3): 460-471.

Wessinger CA, Hileman LC. 2016. Accessibility, constraint, and repetition in adaptive floral evolution. *Dev Biol*.

Wiens JJ. 2011. The causes of species richness patterns across space, time, and clades and the role of "ecological limits". *The Quarterly Review of Biology* **86**(2): 75-96.

Annex 1. Pellissier *et al.* (2016).

Journal of Ecology

British Ecological Society

Journal of Ecology 2016, **104**, 1116–1125 doi: 10.1111/1365-2745.12580

The simultaneous inducibility of phytochemicals related to plant direct and indirect defences against herbivores is stronger at low elevation

Loïc Pellissier[1,2†], Xoaquín Moreira[3†], Holger Danner[4†], Martha Serrano[5†], Nicolas Salamin[5], Nicole M. van Dam[4,6,7] and Sergio Rasmann[8*]

[1]Landscape Ecology, Institute of Terrestrial Ecosystems, ETH Zürich, Zürich, Switzerland; [2]Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland; [3]Biological Mission of Galicia (MBG-CSIC), Apdo. 28, 36080 Pontevedra, Spain; [4]Molecular Interaction Ecology, Institute of Water and Wetland Research (IWWR), Radboud University Nijmegen, PO Box 9010, 6500 GL Nijmegen, The Netherlands; [5]Department of Ecology and Evolution, University of Lausanne, Lausanne CH 1015, Switzerland; [6]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany; [7]Institute of Ecology, Friedrich Schiller University Jena, Dornburger-Str. 159, 07743 Jena, Germany; and [8]Laboratory of Functional Ecology, Institute of Biology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

Summary

1. Ecological theory indicates that warmer and more stable climates should result in stronger biotic interactions. Therefore, plant species growing at lower elevations and experiencing greater herbivore pressure should invest in higher levels of defences than those at higher elevations. Nonetheless, there are a number of studies that have found no effect of elevational gradients on plant defensive traits. Several factors might explain the lack of consistency for the altitude–defence relationships, including (i) the reduction of all defensive traits into one measure of resistance; (ii) not considering plant defence as the simultaneous expression of several defensive traits; and (iii) not considering the relative influence of biotic (e.g. herbivory) and abiotic (e.g. climate and soil conditions) factors associated with the ecological gradient.

2. Here, we present a comprehensive test of the effects of elevation and its associated biotic and abiotic factors on the individual and simultaneous expression of constitutive direct and indirect defences and their inducibility (i.e. expression of defences after herbivore attack). Specifically, we estimated climatic and soil variables and measured herbivore damage and constitutive and jasmonic acid-induced glucosinolate levels in the leaves as a proxy for direct defences, and volatile emission as a proxy for indirect defences in 16 *Cardamine* species naturally growing along the steep elevational gradient of the Alps.

3. Within a phylogenetic comparative framework, we found that species growing at lower elevations invested more in the simultaneous inducibility of both direct and indirect defences, whereas species growing at higher elevations invested more in constitutive direct defences. Although we found strong elevational gradients in herbivory and climatic and soil variables, these biotic and abiotic factors only partially explained elevational patterns in plant defences.

4. *Synthesis.* These results highlight that the complex regulation of multiple defence traits strongly vary across elevational gradients and build towards a better understanding of the multiple mechanisms underlying trait evolution and species interactions along ecological gradients.

Key-words: *Cardamine*, defence induction, elevational gradients, glucosinolates, plant defence syndromes, plant–herbivore interactions, resource availability hypothesis, volatile organic compounds

Annex 2: Cano *et al.* in prep.

Biogeography and Diversification of the New World Thatch Palms (Cryosophileae and Sabaleae: Arecaceae)

Cano, A., Stauffer, F.W., Bacon, C.D., Antonelli, A., Serrano-Serrano, M.L., M. Perret

Abstract

Biodiversity and endemism in the West Indies is the result of a natural island biogeography experiment. The flora of these islands can only be explained by long distance dispersals, probably from the surrounding continents, given that they were never directly connected to the mainland. In order to explore how these colonizations occurred, we selected Sabaleae and Cryosophileae, two sister palm tribes known as the New World Thatch Palms (NWTP), since both have clades in the islands and in the mainland of the Americas. We reconstructed a species level phylogeny of the NWTP and time-calibrated it with available fossil data. We used this phylogeny to estimate the ancestral areas and to track shifts in the diversification rates of these lineages. Our phylogeny confirmed that all the NWTP genera are monophyletic and found the recently described genus Sabinaria as sister to Itaya. The ancestral area of the NWTP was most probably Central-North America, from where Sabaleae colonized the West Indies in two independent events (before the mid Miocene and during the Pliocene, respectively). Cryosophileae dispersed into South America during the Eocene and into the West Indies during the Oligocene. Diversification of this tribe in South America was limited (11 species), associated with East-West Andean vicariances during the late Miocene in Amazonia (Chelyocarpus, Itaya) and Chocó (Sabinaria). Contrastingly, in the West Indies Cryosophileae's diversification was more intense (22 species), resulting in five genera, including its richest genus Coccothrinax, whose radiation coincides with a diversification rate-shift. These results showed that the West Indies were colonized by the NWTP at least three times from the mid Oligocene on, and that these dispersals came from the northern hemisphere. They stress the importance of over-water dispersals to explain current palm distribution patterns, since their long distance dispersal capacity enabled the NWTP to colonize South America before the closure of the Panama Isthmus and the Caribbean islands.

Keywords
Arecaceae; Sabaleae; Cryosophileae, West Indies; South America; Calibrated phylogeny, CISP4, CISP5, matK, PRK, RPB2, diversification, biogeography, BEAST, TreePar, BioGeoBEARS