Floral evolution in Cape *Pelargonium* (Geraniaceae)

Inferring shifts in nectar spur length and pollination

Jens Ringelberg
MSc Thesis
Biosystematics Group, Wageningen University
31-10-2012

Supervisors:
Freek Bakker
Timo van der Niet

Abstract

In this MSc Thesis I investigated the effects of pollination and pollinator-shifts on the speciation of the largely South African genus *Pelargonium*, by using nectar spur length as a proxy for pollination syndrome. During fieldwork in South Africa, nectar spur lengths were collected for over 30 species and 1600 individuals, and another 1300 nectar spur lengths were obtained from herbaria specimen. Based on these measurements it was shown that large levels of spur length variation exist within and between species, populations, and individuals. By mapping nectar spur lengths over species-level phylogenetic trees clear evidence was found for an evolutionary trend towards longer spurs, as well as strong correlations between nectar spur evolution and speciation rates. Furthermore, nectar spurs probably evolve according to the pollinator shift model. These results indicate that nectar spur evolution, and possibly pollinator shifts, play a role in speciation within *Pelargonium*. However, more pollinator observations are needed before any clear conclusions can be drawn on this subject.

Index

Introduction

Plant-pollinator interactions and evolution

Plant speciation

Ever since Darwin's *On the origin of species* (1859) it has been clear that a diversification of some kind between different populations of a species is one of the requirements for actual speciation. The classical view (Schluter 2001) of speciation requires for a population to be divided into two (or more) subpopulations, separated in space (allopatry), this way preventing gene flow from one population to the other. Because the two populations occur in a different habitat, they undergo different selective pressure, and therefore start diverging in morphological or behavioural traits. As an incidental by-product of the adaptation to different habitats, initial reproductive isolation is created. This may be caused by any number of mechanisms, from a different preference in mate-choice (prezygotic) to infertility of hybrids (postzygotic). When, after a period of time, the two subpopulations come into secondary contact, it is often found that hybrids of the two subpopulations have a lower genetic fitness than pure-breeds. If that is the case, reproductive isolation may be actively selected for, again through any form or mechanism. This is known as reinforcement. Once this process is completed, there is no gene flow at all between the two populations, even though they live in the same environment (sympatry). At this point, most biologists consider the two populations to be diverged into different species. As this process requires geographical separation between the two subpopulations, it is known as allopatric speciation. Other known processes are parapatric (with partially overlapping ranges of the two populations) and sympatric speciation, although these methods of speciation are less widely accepted (Thibert-Plante & Hendry 2011).

Instead of categorizing speciation processes based on their geographical component, it is also possible to order these processes based on the mechanisms that drive the evolution of reproductive isolation (Schluter 1998, Orr & Smith 1998, Schluter 2001, Via 2001). With this idea in mind, Schluter (2001) identifies four main modes of speciation: ecological speciation, speciation by divergence under uniform selection, speciation by genetic drift, and polyploid speciation. Ecological speciation states that the initial divergence between (or within) populations is caused by divergent natural selection, and that the final isolation is selected for through reinforcement. Speciation may occur sympatric, allopatric or parapatric, and divergent selection can be based on morphological, sexual or behavioural traits, or any other kind of characteristics. The idea of ecological speciation was developed in the 1940s. Dobzhansky (1951) hypothesized that speciation in *Drosophila* occurs mainly through divergent adaptation to different environmental conditions, and Mayr (1942) understood the importance of some kind of physiological or ecological isolation mechanism. This led to general acceptance of the ecological mode of speciation (Schluter 2001). Speciation by divergence under uniform selection, on the other hand, occurs in different populations undergoing similar selective pressure. If in these populations different incompatible, but equally advantageous mutations arise, populations may divert even though all external conditions are the same. Weinreich *et al.* (2006) showed that many different mutational pathways exist in protein evolution, which could indicate the existence of many different solutions to the same problem. Once again, like in ecological speciation, when these diverged (sub)populations come into secondary contact, reproductive isolation would be actively selected for through the mechanism of reinforcement.

Speciation by genetic drift is based on divergence in traits through the random process of genetic drift. Reproductive isolation is not selected for, at least not until secondary contact takes place, in

which case selection for reinforcement may increase the fitness of both populations. A possible example of speciation by genetic drift could be the amplification of differences in mate preference. Finally, polyploid speciation is most easily distinguishable from the other modes of speciation, because it can be readily diagnosed genetically. It is more common in plants than in animals, but even in plants, it only accounts for 2-4% of all speciation events (Otto & Whitton 2000).
Most evolutionary ecologists agree that in the South African Cape flora, ecological speciation is probably the main mode of speciation (Van der Niet & Johnson 2008). However, considerable debate exists about the exact method of ecological speciation. The two main points of view argue that either pollinator-driven speciation (Van der Niet & Johnson 2008, Johnson 2010, Johnson & Anderson 2010) or soil type-driven speciation (Van der Niet *et al.* 2006, Schnitzler *et al.* 2011) is responsible for the extreme species richness in the South African Cape (Linder 2003).
Van der Niet *et al.* (2006) compared 41 South African sister species pairs belonging to the families Geraniaceae, Iradiaceae and Orchidiaceae. Based on complete pollinator, distribution and edaphic information, they argued that speciation events were most often correlated with soil-type shifts. Pollination diversification mainly was important as the mechanism for reinforcement of reproductive isolation, after initial reproductive isolation was driven by shifts in soil type. They concluded this based on their observation that for sympatric species, pollination shifts were significantly associated with edaphic shifts. On the other hand, in allopatric sister species there was no association between pollinator shifts and edaphic shifts.

In a similar but more extensive study, Van der Niet & Johnson (2009) compared ecological shifts between 188 South African sister species pairs from eight different Cape genera (including *Disa*, *Pelargonium*, and *Satyrium*). They found ecological shifts had occurred in 80% of these pairs, indicating the importance of ecological speciation. Upon comparing the frequency of pollinator, distribution, fire-survival strategy and edaphic shifts, Van der Niet & Johnson concluded that soil-type shifts were in fact quite rare (occurring only in 17% of the 188 sister species pairs). Shifts in pollinator use (33%), distribution (32-62%) and fire-survival strategy (33%) had risen more frequently. Furthermore, more speciation events were accompanied solely by floral rather than vegetative diversification. These findings caused the authors to suggest an important role for pollinator-driven speciation in the South African Cape.

Schnitzler *et al.* (2011) examined 470 species of three of the largest plant families in the Cape: the two Iris genera *Babiana* and *Moraea*, the genus *Protea* and the Podalyrieae. Using near-complete species-level phylogenies, they identified the sister species pairs in these clades. Like Van der Niet & Johnson (2009), they compared the frequency of ecological shifts during speciation events. Remarkably, edaphic shifts were found to be most frequent in *Babiana*, *Moraeae* and *Protea*, and fire-survival strategy shifts were most common in Podalyrieae. Contrary to the conclusions of Van der Niet & Johnson, pollination syndromes showed a high level of phylogenetic conservatism. Based on these results Schnitzler and colleagues reject Van der Niet & Johnson's results, instead stating that pollination-driven speciation plays a relatively minor role in the Cape flora.
A possible explanation for this discrepancy in conclusions may be the fact that Van der Niet & Johnson (2009) included several plant genera with highly specialized pollination systems (*Disa*, *Pelargonium* and *Satyrium* all attract pollinators with their so-called nectar spurs), whereas Schnitzler *et al.* (2011) did not include these species in their analysis (although they did examine species from *Moraea* with specialized pollination systems). Notwithstanding the conclusions of Schnitzler and colleagues, most other publications argue that pollinators have played an important role in the diversification of the Cape flora, if not as drivers of initial reproductive isolation, then at least as method to exert reinforcement upon secondary contact (Johnson 2010, Johnson & Anderson 2010).

Evolution of plant-pollinator interactions

Darwin was the first biologist to propose the idea that plants' floral characteristics are mainly shaped by their interactions with pollinators. Ever since he published his two books about pollination in orchid flowers (1862, 1877), pollination ecology has taken a special position within the evolutionary biology. Darwin was so convinced of the close evolutionary relationship between plants and their pollinators, that he believed that possible pollinators of a certain plant can be deduced based on characteristics of its flowers. This led to his famous prediction of the existence of a hawkmoth with a tongue "capable of extension to a length of between 10 and 11 inches" (1862), after he was being confronted with the Malagasay star orchid (*Angraecum sesquipedale*) with its exceedingly long nectar spurs (up to a length of 12 inches). With the discovery of *Xanthopan morganii praedicta*, the Malagasy form of *X. morganii* with an impressively long tongue (Müller 1873), Darwin was proven right, although in nature this hawkmoth has never been observed pollinating *A. sesquipedale* (Johnson & Anderson 2010).

When Darwin's idea, stating that floral characteristics are shaped by their evolutionary relationship with their pollinators, became more accepted, several pollination biologists started categorizing floral traits based on their respective (putative) pollinator. The first one to do so was Delpino (1873-1874), who categorized flowers according to traits such as shape, scent and colour. Although his ideas were met with considerable criticism (e.g. Müller 1882), the act of categorizing flowers based on their characteristics (used as a proxy for their pollinator) is still quite acceptable in present times. The result of this, the so-called 'pollination syndromes', are sets of floral traits shaped by natural selection imposed by their pollinators (Faegri & Van der Pijl 1979). Pollinators can be either biotic (insects, birds, rodents, bats) or abiotic (water or wind), and floral traits can be any characteristic of a flower: shape, size, colour, scent, nectar qualities, or flowering times, to name but a few. It is commonly thought that plants within the same pollination syndrome share the same pollinator, and that this pollinator can be inferred based on the floral characteristics of the plants.

Even though the use of pollination syndromes is widely accepted, there has been little empirical evidence for the existence of these syndromes. Ollerton *et al.* (2009) scored floral attributes of six plant communities on three continents and collected information about their pollinators. They found that ordination of the plants based on their characteristics did not cluster the plants in distinct syndromes. They also found it impossible to predict the correct pollinator of two third of their sampled species. Based on these results they suggest the pollination syndrome hypothesis should be used with caution, as it may not reflect the natural reality. Johnson (2010) rejects this conclusion, based on the fact that Ollerton *et al.* (2009) only used pollination data at high taxonomic levels, such as Order, whereas much inter-specific variation may exist within these Orders. Johnson advocates the search for syndromes at lower, functional levels, such as pollination guilds. These pollination guilds usually consist of one pollinator that visits several flowers, although exceptions are known. A well-known example is the long-proboscid fly pollination system in *Gladiolus* (Iridaceae), with one or two species of flies pollinating up to 29 plant species (Goldblatt & Manning 1999).

Whether pollination syndromes represent biological reality or not, it seems clear that species richness and diversity are correlated with floral specialization (Van der Niet & Johnson 2012). The most widely accepted explanation of this fact was originally proposed by Darwin (1876), who stated that floral adaptation to pollinators resulted in convergent evolution among species that share a pollinator. Ultimately this would lead to diversification between plants within genera. This theory was further elaborated by Grant and Stebbins (Grant 1949, Grant & Grant 1965, Stebbins 1970). Therefore, the model of pollinator-driven plant diversification (in angiosperms) is now known as

the Grant-Stebbins model. A highly important aspect of pollination-driven clade-proliferation are pollinator-switches (Johnson 2010). During the development of their model, Grant (1949) and Stebbins (1970) placed much emphasis on the importance of pollinator-switches. Many studies have confirmed their hypothesis by demonstrating that switches in pollinator use have occurred frequently during clade radiation, both globally (Givnish & Sytsma 1997, Weller & Sakai 1999) as in South Africa (in almost every speciation event in *Disa*, a pollinator switch took place (Johnson *et al.* 1998), and in *Gladiolus* and *Babiana* (Iridaceae), Goldblatt & Manning (2006) estimated at least one switch for every five to six species).
Johnson (2010) identified five different modes of diversification driven by pollinators: pollination system shifts, where local shifts in pollinators create a geographical mosaic with much variation between different populations of the same species; divergent use of the same pollinator, such as placement of pollen on different parts of the pollinator's body, leading to diversification between (or even within) a population; coevolution, leading to an evolutionary arms race between plant and pollinator; trait tracking, in which different members of a pollination guild are forced to closely resemble the most prominent plants within the guild; and mimicry of different model flowers, where a non-rewarding species mimics two different model flowers, which will cause diversification. Various levels of evidence exist for all these modes of diversification.

Nevertheless, Darwin's explanation of the correlation between floral specialization and species diversity is disputed by some authors. Armbruster & Muchhala (2007) reviewed the existing literature and tested different hypotheses in four study systems. They reached the conclusion that in most cases floral specialization is the *effect*, rather than the *cause*, of species diversity. Whereas Darwin's theory states that different pollination systems cause floral specialization and ultimately lead to species diversity, Armbruster & Muchhala argue that local species diversity drives the evolution of specialized pollination, the so-called character displacement hypothesis. However, the authors make an exception for specialized pollination systems: they state that it is highly likely that floral specialization may promote species diversity in these systems, either through initial reproductive isolation or through reinforcement upon secondary contact.

Nectar spurs

One form of a specialized pollination system is a nectar spur. Nectar spurs consists of a long tube at the base of a flower, either adnate along the pedicel or free behind the petals. Nectar is secreted at the base of the tube, which prompts the pollinator (usually insects, but birds or even bats are possible as well) to extend its tongue/proboscis into the spur. If spur and proboscis/tongue length are closely matched, the pollinator may only reach the nectar when it presses its body against the reproductive organs of the flower, this way either acquiring or releasing previously accumulated pollen (Hodges 1997). As pollinators preferably only visit plants with spurs that are the same length (or smaller) as their proboscis (in order to be able to reach the nectar at the base of the spur) and plants prefer to attract pollinators with proboscises that are similar in length (or smaller) to their spurs (to promote the dispersal of pollen, a pollinator needs to press its body against the fertile parts of the plant), there exists a close evolutionary relationship between pollinators and plants in this specialized pollination system (Hodges & Arnold 1995, De Wet *et al.* 2008, Johnson *et al.* 2010).

Nectar spurs evolved several times in history, in different plant genera and families. Flowers of families such as the Balsaminaceae, Fumeraceae, and Lentibulariaceae and genera like *Aquilegia*, *Disa*, and *Pelargonium* all have nectar spurs (Hodges 1997). In a comparison of six taxa with spurs with their non-spurred sister taxa, Hodges & Arnold (1995) found that five of these taxa underwent significant species diversification. Only *Pelargonium* contains less species than its sister genera, all other taxa were considerably more speciose than their sisters. This caused Hodges & Arnold (1995)

to strongly support the hypothesis that nectar spurs represent a key innovation, a morphological or behavioural trait that supports species diversification by offering the ability to rapidly speciate after environmental change (Liem 1973) or by opening new 'adaptive zones' (Simpson 1953). In several genera, attempts have been made to identify the genetic factors underlying nectar spur development; for example in *Aquilegia* (Ranunculaceae, Kramer & Hodges 2010), *Platanthera* (Orchidaceae, Little *et* al. 2005), and *Antirrhinum* (Scrophulariaceae, Box *et al.* 2011). Information about the genetic basis of nectar spur development in *Pelargonium* is not available, and in this thesis I will not deal with this aspect of nectar spur evolution.

The general agreement within the scientific community is that nectar spur length and pollinator proboscis/tongue length are closely matched because of their evolutionary relationship. Most pollination biologists share the view of Darwin (1862), who stated that as pollinators prefer to visit plants with spurs equal to or shorter than their proboscids, and plants aim to attract pollinators with proboscids equal to or shorter than their spurs, plants and pollinators are locked in an evolutionary arms race: both continuously try to outgrow their 'opponent'. This may eventually lead to species with rather absurd characteristics, such as *Angraecum sesquipedale* and *Xanthopan morganii praedicta*. Nowadays, the evolutionary arms race hypothesis is considered to be encompassed by the more general evolutionary force of coevolution (Pauw *et al.* 2008). Many authors agree that this model of coevolution is most likely to drive the diversification of both pollinator and plant species (Johnson & Steiner 1996, Ennos 2008, Pauw *et al.* 2008, Rodríguez-Gironés & Llandres 2008, Johnson & Anderson 2010).

However, some studies point in a different direction. Whittall & Hodges (2007) hypothesized that the relationship between pollinators and columbine flowers (*Aquilegia*) is only one-sided: pollinators may drive plant spur diversification, but not vice versa. They state that as the pollinators of *Aquilegia* are much older than the *Aquilegia* genus itself, pollinator proboscis lengths were already fixed before coming into contact with *Aquilegia*. Changes in spur length would then be caused by the switching of plants to a different pollinator with a different proboscis length. Changes in spur length would occur rapidly, but with long intervals of no change, according to this model, instead of gradually, but continuously, as during the evolutionary arms race. This theory is known as the pollinator shift model, and at least in the North-American genus *Aquilegia* considerable evidence was found for its occurrence (Whittall & Hodges 2007 and 2008).

In recent studies of the South African flora, these two models (the evolutionary arms race model and the pollinator shift model) were shown to be not mutually exclusive. Janzen (1980) coined the term 'diffuse coevolution' for the situation in which several populations, belonging to different species, participate in reciprocal selection with a different population of species. Pauw *et al.* (2008) showed that this is likely to be the case for several plant-pollinator communities in South Africa, in which plants help shape their pollinators proboscids, and vice versa, through coevolution. However, as rare or new plant species within such communities are not common enough to execute selective pressure on their pollinators, these plants are only engaged in one-sided evolution: they have to start resembling the other plants within the community to attract the attention of the pollinators. This way both coevolution (the evolutionary arms race model) and one-sided evolution (the pollinator shift model) help shape the communities of plants and pollinators (Pauw *et al.* 2008, Johnson & Anderson 2010).

Not only does considerable debate exist about the mechanisms of spur length evolution, also about the existence (and if so, direction) of trends in spur length evolution the opinions are divided. Whittall & Hodges (2007) found an evolutionary trend in *Aquilegia* from pollination by bumble-bee to pollination by hummingbird to finally pollination by hawkmoth. This means that the main

pollinators' proboscis/tongue length increased over time, and the nectar spurs evolved correspondingly. Bakker *et al.* (2005) found a similar trend from short to longer spurs and proboscises in (some clades) in the genus *Pelargonium*. This is often thought to be the main trend within nectar spur evolution, but considerable evidence exists of the opposite (Johnson & Anderson 2010). On a macro-evolutionary scale, Micheneau *et al.* (2008, 2009) reported shifts from hawkmoth to bird pollination in *Angraceum*, with a corresponding decrease in nectar spur length. And Bloch & Erhardt (2008) provided micro-evolutionary evidence that shorter spurs may be selected for when pollinators have short proboscids. With different models, Rodríguez-Gironés & Santamaría (2007, 2010) and Rodríguez-Gironés & Llandres (2008) showed that nectar spur elongation and shortening may evolve under the influence of competition for nectar and pollinators. According to their models, floral and nectar spur divergence, rather than spur elongation, may be selected for.

The Cape Floristic Region

Found at the most southern tip of the African continent, the South African cape has since long been known for its hyperdiverse flora. It is recognised as one of the six Floristic Kingdoms of the world (Takhtajan 1986) based on its extraordinarily high levels of vascular plant species richness and endemism and is considered to be a model system for plant evolution (Linder 2003). As the Cape flora is so unlike the flora in the rest of the continent, the area is known as the Cape Floristic Region (CFR; Goldblatt 1978) or the Greater Cape Floristic Region (GCFR; Born *et al.* 2006). The CFR contains the extremely diverse fynbos biome, whereas the GCFR also contains the more arid succulent karoo biome. There are considerable differences between these two biomes: the fynbos flora consists mostly of heathy evergreen shrubs, whereas the succulent karoo flora is characterised by life forms who show a whole range of mechanisms to tolerate the extreme drought (Verboom *et al.* 2009). However, there are many resemblances in the floras of these two biomes as well, which may indicate a partially shared evolutionary history. The fact that both biomes share a winter-rainfall regime, unlike the rest of the African continent, may also explain their similarities (Linder 2003).
The CFR was first surveyed by Thunberg, a student of Linnaeus, who wrote an extensive book about the Cape flora: *Flora Capensis* (1813). In the following decennia, the extreme plant diversity was further examined by the commercial plant collectors Drège, Ecklon and Zehyer and by the botanist Bolus (who founded the Bolus Herbarium in Cape Town) (Linder 2003). In the 20th century, important monographs about the Cape flora were written (Marloth 1908, Goldblatt 1978), numerous collecting trips were made and extensive catalogues were composed (Bond & Goldblatt 1984). As a result of this extensive botanic activity over the last 250 years, the Cape flora is rather well documented (Linder 2003).

Species endemism and species richness

Endemism levels in the CFR range between 68.8% (at specific level) and 16.2% (at generic level) (Goldblatt & Manning 2000), which is extremely high for continental areas. In fact, based on its endemism levels, the South African Cape can be compared to islands such as Hawaii, New Zealand and Madagascar (Linder 2003). High levels of species endemism on islands can be explained by the geographical isolation of these islands, and in the Cape a similar explanation may be used. After all, the Cape is surrounded by oceans on three sides and the only link to the tropical African flora in the north is blocked by large arid areas and deserts. Furthermore, apart from being geographically isolated from the rest of the continent, ecological isolation, based on the completely different soil and climate in the CFR, contributes to the high levels of species endemism as well.

Species richness levels of the CFR, however, cannot be compared with richness levels of islands. Based on its species richness, the area ranks with tropical rainforests, such as those of Panama, the Philippines or Brazil, rather than with other Mediterranean areas or islands (Linder 2003). Given its relatively small size, a much lower species richness would be expected. High species richness can either be explained by high levels of speciation or low levels of extinction (Barraclough 2006), with the first explanation receiving the most attention (e.g. Linder & Hardy 2004, Verboom *et al.* 2008, Carlson *et al.* 2010, Rymer *et al.* 2010). Most authors cite ecological speciation (Schluter 2001) as the most important speciation method in the CFR, but, as was mentioned before, there is still considerable debate about the exact drivers of speciation, such as soil-type shifts, pollinator specialisation and fire-survival strategies (Van der Niet *et al.* 2006, Van der Niet & Johnson 2009, Schnitzler *et al.* 2011). Recent climatic change may have had a considerable impact on speciation rates in the Cape as well (deMenocal 2004, Warren *et al.* 2011).

The South African Big Genera Group

A large part of the species richness in the CFR is accounted for by a relatively small amount of clades (Linder 2003). The South African Big Genera Group (SABiGG, http://www.reading.ac.uk/AcaDepts/ap/SABGG/publish/index.htm) was a group of scientists working on making phylogenetic reconstructions of the main families and genera of the Cape Floristic Region. The group made a list of the twenty largest Cape genera, which includes genera such as *Disa* (Orchidaceae), *Erica* (Ericaceae), *Moraea* (Iridaceae), *Oxalis* (Oxalidaeceae), *Pelargonium* (Geraniaceae) and *Restio* (Restionaceae). Many SABiGG members have published studies and phylogenies about these genera (e.g. Oberlander *et al.* 2002, Bakker *et al.* 2004, Verboom *et al.* 2004, Eldenas & Linder 2010).

Apart from endorsing individual studies, SABiGG also published metastudies as a group. In the first metastudy (Verboom *et al.* 2009), the origin and diversification of the two main GCFR biomes, fynbos and succulent karoo, were examined. By comparing several Cape-endemic groups of plants, it was found that succulent karoo probably radiated quite recently, less than 17.5 million years ago. Fynbos, on the other hand, is probably much older, with some lineages originating at least 23 million years ago, although some other lineages were much younger. Furthermore it was found that in both fynbos and succulent karoo considerable speciation has occurred since the origin of these biomes.

In the second metastudy, Warren *et al.* (2011) examined the effects of the recent changes in rainfall seasonality in the South African Cape. Plant species generally have two defence mechanisms to such seasonal climatic changes: either they move away (distributional shift) or they change their flowering times (phenological shift). About half the Cape clades examined underwent distributional shifts, the other half underwent phenological shifts. Of the extant Cape species, 14-41% originate from lineages that shifted in distribution, and 14-41% shifted in phenology. This suggests an important role for recent climatic changes in the making and diversification of the flora of the South African Cape. However, without a doubt there is still much to discover about this fascinating region.

Pelargonium

The genus *Pelargonium* L'Hér, one of the main genera in the Geraniaceae, consists of 280 species, most of which occur in South Africa. About 200 of the *Pelargonium* species occur in the Winter Rainfall Region of the CFR, making it the third largest angiosperm CFR genus (Goldblatt & Manning 2000). Financially speaking it is an important genus as well, as cultivated *Pelargonium*

species (often wrongly labelled as Geranium) are found in almost every garden and house in the world. This is reflected by the rich horticultural history of the genus. As Cape Town has been an important seaport for centuries, the first *Pelargonium* species were taken back to Europe as early as 1690, and have been thoroughly cultivated and hybridized since then, resulting in species such as regal geranium (*Pelargonium x domesticum*) and zonal geranium (*Pelargonium x hortorum*) (Loehrlein & Craig 2001). For a more complete overview of the horticultural history of *Pelargonium*, see the booklist of the Geraniaceae Group (http://www.geraniaceae-group.org/booklist.html).

Most *Pelargonium* species are described in '*Pelargoniums* of South Africa' by Van der Walt (1977) and Van der Walt & Vorster (1981, 1988). Precise descriptions of both sections and individual species of *Pelargonium* by botanists and biologists such as Albers (e.g. Albers *et al.* 1991), Gibby (e.g. Gibby *et al.* 1996) and Marais (e.g. Marais 1996) further increased the scientific knowledge about this genus. As a result, extensive information is available about the various forms in which *Pelargonium* species occur.

Variation

Pelargonium is known for its extensive vegetative and floral variation. The genus includes life forms such as stem succulents, woody and succulent (sub)shrubs), geophytes, and herbaceous annuals. These various forms probably reflect the different conditions under which the plants occur. Some aspects of the plant, such as stem succulence and the formation of tubers, appear to have evolved separately on several occasions (Bakker *et al.* 2005). Recently, Jones *et al.* (2009) showed that individual leaf elements varied in their transformation rates. Leaf elements such as leaf base, apex and overall outline had a relatively slow transformation rate, whereas elements such as the extent of lamina lobbing and functional leaf area evolved much faster. Jones and colleagues argued that some of this variation was caused by being under different natural selection, but could not entirely rule out the possibility that some of variation was non-adaptive. *Pelargonium* has a wide range of floral variation as well, which probably reflects adaptation to different pollinators (Struck 1997).

The wide range of vegetative and floral variation in *Pelargonium* is reflected by a considerable genomic instability of the genus. Guisinger *et al.* (2008) showed that plastid genomes of *Pelargonium* and other members of the Geraniaceae are relatively large and extensively rearranged. In fact, rates of amino acid change were so high, that they pointed in the direction of positive or relaxed plastid selection, which is rather hard to explain given our current understanding of plastid evolution in photosynthetic plants. Exceptional substitution rate accelerations were also found in mitochondrial DNA in the Geraniaceae (Bakker *et al.* 2006). In both studies, the authors offer explanations such as altered gene expression and affected proofreading accuracy control and repair to account for the wide genomic instability found in the genus. Furthermore, there is a wide range of basic chromosome numbers throughout *Pelargonium* (Bakker *et al.* 2005).

Phylogeny

Pelargonium can be divided into two main clades: one with small chromosomes (<1,5 µm) and one with large chromosomes (1,5-3,0 µm). The group with small chromosomes consists of up to 80% of the species of the whole genus (Bakker *et al.* 2000). The genus consists of 6 subclades and 16 recognised sections. Ordered by main clade and subclade, these are the 15 sections:

- large chromosomes clade:
    - clade C1: sections *Jenkinsonia*, *Chorisma*, and *Myrrhidium*

    - clade C2: sections *Subsucculentia* and *Ciconium*
- small chromosomes clade:
    - clade B: sections *Peristera* and *Reniformia*
    - clade A1: sections *Campylia* and *Pelargonium*
    - clade A2POC (A2a in Bakker *et al.* 2005): sections *Otidia*, *Cortusina*, and *Polyactium*
    - clade A2HLM (A2b in Bakker *et al.* 2005): sections *Magnistipulacea*, *Ligularia*, and *Hoarea*.

In this thesis, a new phylogenetic tree will be made for *Pelargonium* (see chapter 2).

Pollination

Notwithstanding two or three exceptions, all *Pelargonium* species have a tube filled with nectar adnate along the pedicel. This structure is usually described with the term 'hypanthium' (e.g., see Van der Walt 1977). In this thesis, I will treat this structure as nectar spur (as was done by Hodges 1997).

In a genus-wide analysis of floral characteristics, overview of the existing scientific literature, and field observations, Struck (1997) inferred the main pollinators of *Pelargonium*. It was found that 60% of species in the genus are pollinated by bees, 25% by long-proboscid hovering flies, 7% by moths, 2-4% by butterflies and 1% by birds. Struck hypothesized that nectar spur length, together with several other characters such as colour and shape of the flower and number of flowers per inflorescence, are important pollination factors, shaped by the evolutionary relationships between plants and their pollinators. Similar conclusions were reached by Marais (1999) in an analysis of pollination in section *Hoarea*.

Furthermore, Struck (1997) deemed it highly likely that there had been several pollinator shifts during the evolutionary history of *Pelargonium*: “Moreover, it seems likely that shifts in pollination systems occurred rapidly within *Pelargonium,* as can be deduced from the deviations in floral characters and in pollination syndromes which repeatedly occur in the intraspecific level, e.g., in *P. antidysentericum*.” Indeed, pollinator shifts within and between species have been demonstrated in several studies. Van der Niet & Bakker (unpublished) showed there had been pollinator shifts between the closely related species *P. candicans*, *P. longicaule*, and *P. myrrhifolium*, and even a shift from pollination by flies to pollination by bees within *P. longicaule*. The authors stated that these shifts were best explained by the Grant-Stebbins model of pollinator-driven diversification. In *P. reniforme*, De Wet *et al.* (2008) showed there are two different subpopulations: one pollinated by pollinators with short proboscises, and the other by pollinators with long proboscises. The nectar spurs of the two subpopulations evolved correspondingly to fit the proboscis of the pollinators. De Wet and colleagues argued that a shift in pollinator use could explain these differences. Finally, in *P. alternans*, a similar pattern was distinguished: one subpopulation of the species adapted to a different pollinator, and changed its floral characteristics accordingly (Becker & Albers, unpublished). It seems likely to assume that pollinator-driven diversification plays an important role in *Pelargonium*, and that there have been several more pollinator shifts in *Pelargonium*.

Hypotheses and premises

The following hypotheses have been formulated for this project:

1. *Pelargonium* nectar spur length is constant at the population level.

2. Pollinator-switches and clade-proliferation are linked in clades A2HLM and A2POC.

3. In the evolution of nectar spurs in *Pelargonium*, there was a trend towards longer spurs.

In order to address these hypotheses, I have to make several assumptions. These three premises were identified:

1. The spur length of a *Pelargonium* species is a 'one-dimensional proxy' for its pollinator syndrome.

2. Phylogenetic trees accurately portray the evolutionary history of the genus.

3. During its evolution *Pelargonium* underwent several pollinator-switches.

The hypothesis will be addressed more elaborately in the following chapter. The first chapter of this thesis will deal with hypothesis number one, while hypothesis two and three will be tested in chapter two. The final chapter, chapter three, will mainly deal with the first premise, which states that nectar spur length is a proxy for a pollination syndrome.

Chapter 1:
Nectar spur length distribution

Introduction

This chapter deals with the first hypothesis of this thesis, which states: '*Pelargonium* nectar spur length is constant at the population level'. This hypothesis seems intuitively right; if there would be much intraspecific variation, individuals of the same species would attract different pollinators, which would cause reproductive isolation (and could be the start of speciation). In most of the scientific literature about nectar spurs (e.g., see Hodges & Arnold 1995, Rodríguez-Gironés & Llandres 2008), it is assumed that intraspecific variation in nectar spur length must be low, and interspecific variation high, in order for different plant species to attract different pollinators. This point of view is supported by the works of Van der Walt (1977) and Van der Walt & Vorster (1981, 1988), which report little variation in spur length of South African Pelargonium species.

However, some objections can be made against this hypothesis. First of all, spur lengths in the scientific literature are often based on surprisingly few measurements. For example, Whittall & Hodges (2007) measured spur lengths of 10-52 individuals in 1-5 populations per *Aquilegia* species. This may be perfectly acceptable, but it is possible that much spur length variation remains hidden this way. Furthermore, several studies have made clear that much variation in spur length exists between populations of the same Pelargonium species. For example, the variation in *P. longicaule* was documented by Pauw *et al.* (2008) and De Wet *et al.* (2008) noticed the variation in spur length between different populations of *P. reniforme*. Based on these studies, it certainly is a possibility that spur length variation exists within and between Pelargonium populations.

In order to test this hypothesis, I travelled to South Africa to measure nectar spur lengths of a range of *Pelargonium* species, both in natural populations in the field, as well as in South African herbaria. Furthermore, I collected spur length information from different sources in the scientific literature. The resulting dataset comprises of nectar spur lengths of over 180 species, and was statistically analysed to gain further insight in spur length distribution and the level of interspecific and intraspecific variation.

Materials and Methods

Field protocol

To make sure the collecting and measuring of nectar spurs in the field occurred in an organised and consistent way, a field protocol was developed. By adhering to this protocol it was made sure that nectar spurs collected on different days or in different locations could easily be compared. It also made sure there would be no bias, neither in the spurs collected, nor in the way the spurs were measured.

The complete fieldwork protocol consists of the following points:

- of each population of *Pelargonium* species that was encountered, one nectar spur per individual was collected;
- collected nectar spurs were photographed in a petri dish (for protection against wind) against a background of millimetre-scaled paper (see figure 1.1), to allow later measuring of the nectar spur;
- as the styles of *Pelargonium* flowers elongate after fertilisation (hence the name *Pelargonium*, which means storksbill) only nectar spurs of non-pollinated flowers were collected;
- to avoid measuring nectar spurs that have changed in length due to a lack of sunlight (E.M. Marais, pers. comm.), only well-exposed nectar spur were collected;
- to avoid measuring nectar spurs that were not yet fully grown, only fertile flowers were collected from fully grown, adult individuals;
- no distinction was made between male and female flowers;
- subject to practicality, altitude level and GPS coordinates were recorded for each population.

Figure 1.1. A nectar spur of *Pelargonium alchemilloides* collected during the fieldwork. The background of millimetre-scaled paper allows easy measuring of length of the nectar spur.

As different *Pelargonium* species have different habitats and different distributions, I have tried to visit as many different fieldwork locations as possible. However, due to constraints considering both time and money, only locations in the vicinity of Cape Town were visited.

Herbarium protocol

As South Africa's flora has been extensively collected and documented by botanists for more than a century, many herbaria have considerable collections of *Pelargonium* species. And as it is impossible to measure every *Pelargonium* species in the field due to constraints on money and time, herbarium collections provide a useful alternative. During this project, I visited four herbaria: the Bolus Herbarium (BOL, Cape Town, South Africa), the Compton Herbarium (NBG, located in the Kirstenbosch National Botanical Garden, Cape Town, South Africa), the Stellenbosch University Herbarium (STEU, Stellenbosch, South Africa), and the Herbarium Vadense (WAG, Wageningen, The Netherlands). For similar reasons as during the fieldwork, a herbarium protocol was developed.

The herbarium protocol consists of the following points:

- one nectar spur per collected specimen was measured;
- only non-pollinated nectar spurs were measured;
- measuring occurred using a pair of callipers;
- if possible, only specimen directly collected from the wild were measured - although in some cases only specimen grown in botanical gardens were available;
- species name, collector name and date, and, if present, GPS coordinates were recorded.

Nectar spur transformations

One of the problems of measuring plant characteristics from herbaria specimen, is that these specimen may have changed in shape, form or colour since they were collected in the field. To account for the possibility that nectar spurs may have changed in length in the herbarium, an experiment was performed. Flowers of two plants (25 flowers of *Pelargonium sidoides* and 20 flowers of *Pelargonium zonale*) were collected, measured, weighed, placed in a plant press (with varying amounts of pressure), and put in a stove for 22 hours (65 degrees Celsius). The following day, the flowers were measured and weighed again, and the differences were noted. A Student's t-test for paired samples (Moore & McCabe 2009) was performed in R (R Development Core Team 2012) to verify whether there had been a significant change in nectar spur length or flower weight.

Spur length measurements

The result of my fieldwork in South Africa consisted of 1685 standardized photographic recordings of nectar spurs, as will be seen further in this chapter. To analyse these results, and obtain the exact nectar spur lengths, I used two computer programs: tpsUtil (Rohlf 2010b) and tpsDig (Rohlf 2010a).

tpsUtil is simply a utility program, used to transform .JPG-files to .TPS-files, the file extenstion used by tpsDig. tpsDig can be used to measure lengths, curves or surfaces of objects depicted on a photograph, as long as the scale of the photograph is known. As I photographed all the nectar spurs on millimetre-scaled paper, the collected nectar spurs could be measured in millimetres.

The procedure for measuring a (curved) nectar spur can be seen in figures 1.2 and 1.3.

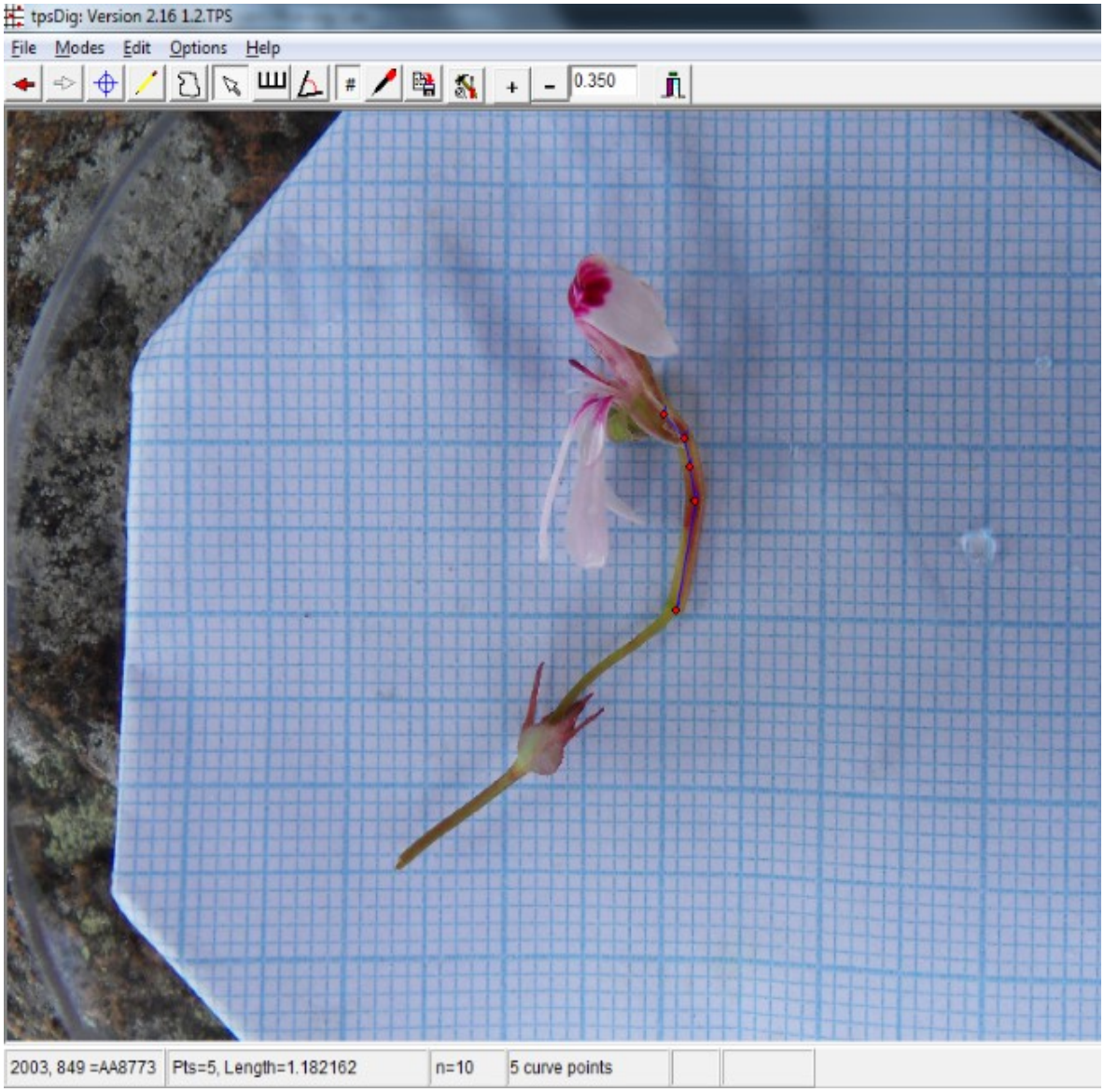

Figure 1.2. First step of measuring the length of a nectar spur using tpsDig. On the photograph of the flower, the curve and length of the nectar spur are identified by locating several marking points (red points on blue line) on the nectar spur.

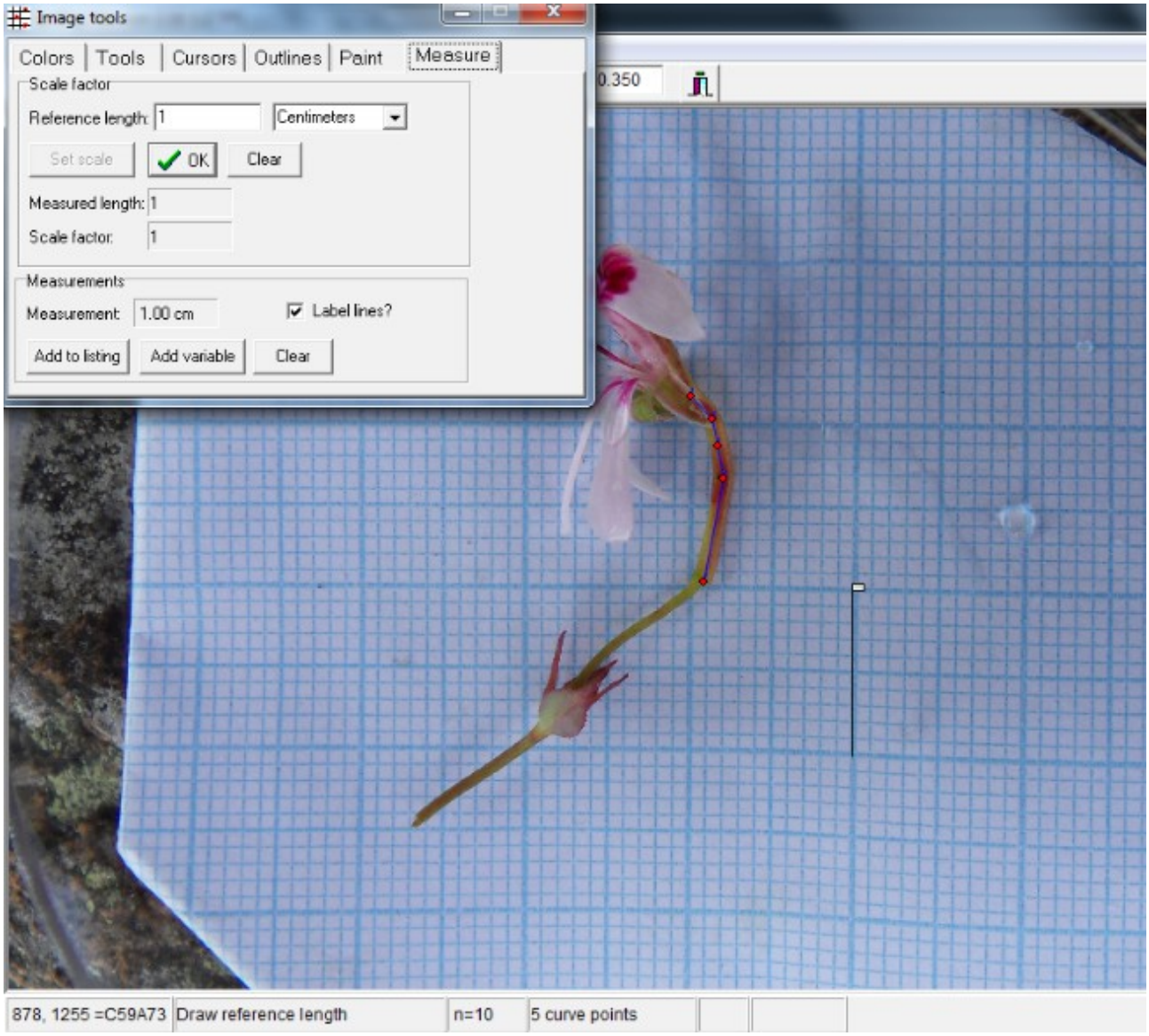

Figure 1.3. Second step of measuring the length of a nectar spur using tpsDig. Once the nectar spurs has been marked, the scale of the picture can be established by determining a reference length (the black line with the small white rectangle). Here, the reference length is one centimetre, as can be seen in the "Image tools"-window. Now the length of the nectar spur can be recorded.

Statistics

The combined dataset with different measures of nectar spur length was analysed using the standard packages stats, base, and graphics in R. Average minimum, median, maximum, and range values of nectar spur length were calculated, and the distributions of these measurements were plotted in histograms, using the function "hist". Furthermore, for the field measurements, boxplots of the spur lengths of the different populations were obtained using the function "boxplot".

To test for correlation between these different measures, Pearson's Product-Moment Correlation Coefficient (Moore & McCabe 2009) was calculated using the function "cor.test". Three different sources of nectar spur lengths were distinguished: spur lengths obtained from wild populations in the field, spur lengths obtained from herbaria, and spur lengths from the final dataset, consisting of a combination of measurements from the field, the herbaria and the scientific literature (see further

in this chapter). The field measurements were subdivided into two groups: measurements grouped per species, combining all the populations into one pool; and measurements grouped per population.

Correlation coefficients and significance levels were calculated for any combination of minimum, median, maximum, and range of nectar spur length, for any of the four different groups, resulting in 24 different correlation coefficients.

Results

Fieldwork

During my time in South Africa, I was able to go on 19 fieldwork expeditions, during which I collected 1685 (standardized photographic recordings of) nectar spurs, belonging to 90 populations, 35 different species (of which 30 have been identified), and 10 different sections. Nectar spurs were collected both in Cape Town and in places such as Kogelberg, Cape Point, the Cederbergen and Elandsberg Reserve. A variety of different soil types and habitats was sampled, including fynbos,mountainsides, coastal areas, and karoo. For a full overview of the fieldwork results see table 1.1, and for all the fieldwork locations, see figure 1.4.

In addition to the usual practice of sampling one nectar spur per individual, to measure the variation in nectar spur length within and between populations, I also wanted to gain insight in the variation in spur length within individuals. Therefore I collected 10 flowers per individual, for 10 individuals of *Pelargonium cucullatum*, all belonging to the same population, and analysed the differences.

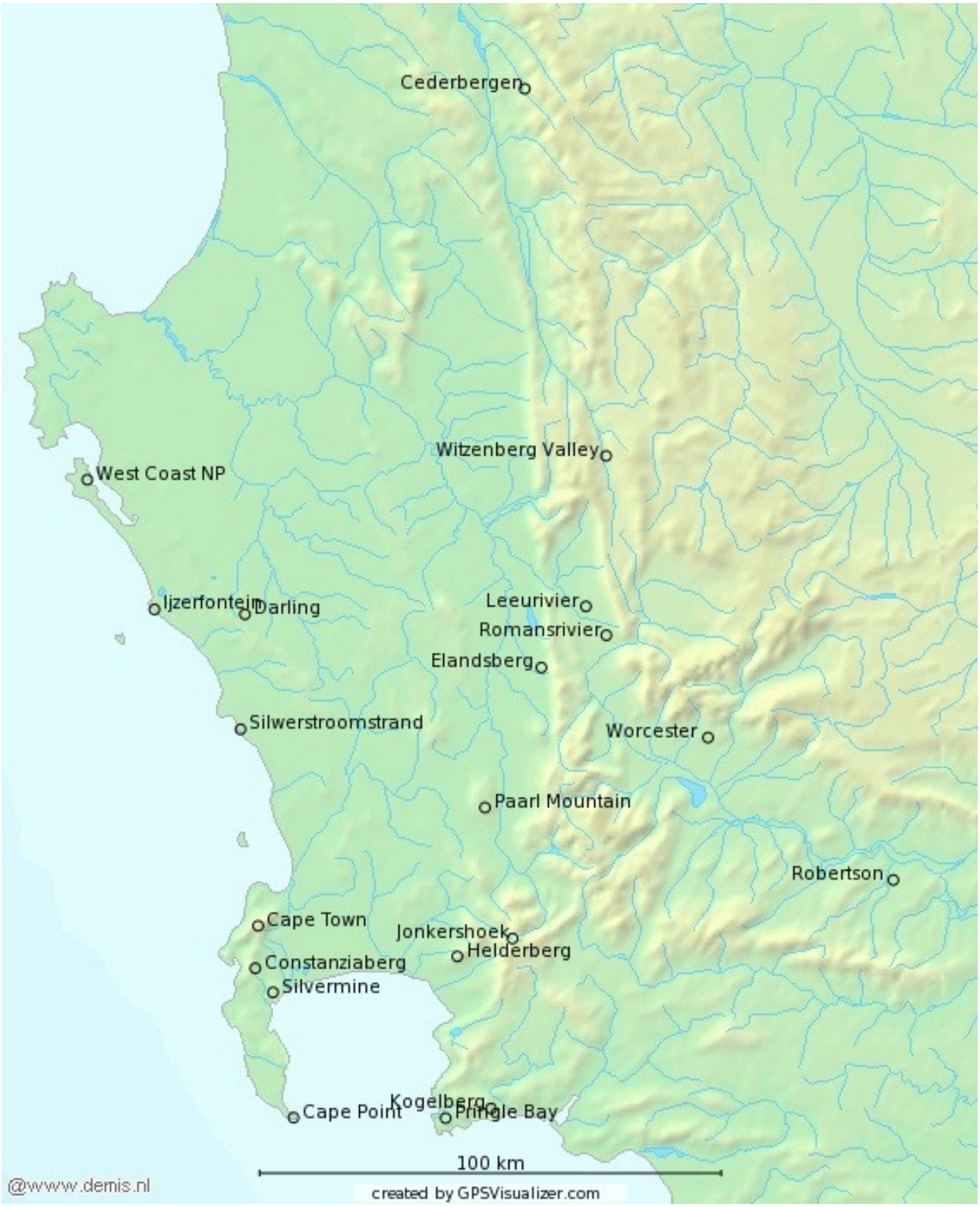

Figure 1.4. Main locations in South Africa where nectar spurs have been collected and phtographed. For a full list, including GPS coordinates and species names, see table 1.1. This map was created using GPSVisualizer (www.gpsvisualizer.com, created by Adam Schneider).

Date	Location	Degrees South	Minutes	Seconds	Degrees East	Minutes	Seconds	Altitude (m)	Species	Section	Amount
31-10-11	**Helderberg Nature Reserve**	34	2	17	18	51	56	913	P. alchemilloides	Ciconium	12
									P. myrrhifolium	Myrrhidium	10
		34	3	10	18	52	27	241	P. alchemilloides	Ciconium	34
									P. myrrhifolium	Myrrhidium	5
									P. triste	Polyactium	6
		34	3	18	18	52	31	223	P. alchemilloides	Ciconium	49
		34	3	17	18	52	28	202	P. triste	Polyactium	5
02-11-11	**Kirstenbosch**	33	58	42	18	24	8	677	P. myrrhifolium	Myrrhidium	32
		33	59	41.81	18	24	47.04	648	P. longicaule	Myrrhidium	11
		33	59	39.77	18	25	8.3	438	P. cucullatum	Pelargonium	9
		33	59	39.77	18	25	8.3	438	P. longicaule	Myrrhidium	24
		33	59	39.6	18	25	10.94	402	P. longicaule	Myrrhidium	24
		33	59	39.6	18	25	10.94	402	P. cucullatum	Pelargonium	13
03-11-11	**Rondebosch Common**	33	57	9	18	29	5		P. triste	Polyactium	37
		33	57	7	18	29	1		P. alchemilloides	Ciconium	25
									P. myrrhifolium	Myrrhidium	29
		33	59	33	18	29	8		P. grossularioides	Peristera	26
									P. myrrhifolium	Myrrhidium	12
									P. triste	Polyactium	14
04-11-11	**Signal Hill and Lion's Head**	33	55	9.08	18	23	57	292	P. lobatum	Polyactium	30
		33	55	3	18	24	5.3	315	P. alchemilloides	Ciconium	53
		33	55	5	18	23	59.39	296	P. auritum	Hoarea	1
		33	55	0.9	18	24	20	308	P. myrrhifolium	Myrrhidium	28
		33	56	11	18	23	15		P. elongatum	Ciconium	13
		33	57	3.44	18	24	10.78	483	P. longicaule	Myrrhidium	15
07-11-11	**Silvermine**	34	6	29.96	18	26	12.64	390	P. longicaule	Myrrhidium	42
		34	6	48.3	18	26	19.68	494	P. alchemilloides	Ciconium	5
		34	3	40.32	18	23	44.12	632	P. zonale	Pelargonium	4
10-11-11	**Kogelberg**	33	57	14.7	18	28	57	16	P. auritum	Hoarea	44
		34	19	21.9	19	6	27.1	55	P. longifolium	Hoarea	47
		34	20	37	19	2	9.8	16	– unknown –		23
		34	18	55.24	18	57	37.22		P. pinnatum	Hoarea	8
		34	20	6.6	18	56	37.8	331	P. setulosum	Pelargonium	30
									P. setulosum	Pelargonium	29
									P. betulinem		16
		34	21	14.7	18	50	17.7	32	P. proliferum	Hoarea	8
12-11-11	**Cederbergen**	32	20	33.2	19	1	19.9	400	P. coronopifolium	Campylia	14
									P. scabrum	Pelargonium	28
12/11/11(?)	**Robertson**	33	53	19.24	19	52	45.48	182	– unknown –	Hoarea	21
		33	41	40.9	19	35	44.9	293	P. trifidum	Jenkinsonia	11
14-11-11	**Rhodes Memorial**	33	57	5.83	18	27	23.88	209	– unknown –		20
		33	57	4.19	18	27	22.09	222	– unknown –		36
		33	56	58.51	18	27	25.68	210	P. vitifolium	Pelargonium	10
									P. senecioides	Jenkinsonia	18
		33	56	34.67	18	26	45.02	220	P. graveolens	Pelargonium	29
		33	57	0	18	26	29	597	P. tabulare	Glaucophyllum	12
15-11-11	**Cape Point**	34	21	11.1	18	29	5.2		P. longicaule	Myrrhidium	17
		34	17	37.6	18	25	53.9		P. grossularioides	Peristera	12
		34	18	5.7	18	25	14.2	86	P. longifolium	Hoarea	20
16-11-11	**Meadowridge Common**								P. triste	Polyactium	28
17-11-11	**West Coast National Park**	33	35	38.7	18	21	41.1		P. gibbosum	Polyactium	38
		33	14	40.8	18	11	31	96	P. senecioides	Jenkinsonia	13
		33	14	38.7	18	11	25.9	96	P. longicaule	Myrrhidium	31
		33	8	52.2	18	0	3	6	P. fulgidum	Polyactium	32
		33	6	23.2	18	0	18.4	171	P. carnosum	Otidia	11
		33	21	35.4	18	9	40.7	59	P. lobatum	Polyactium	3
		33	22	10.2	18	22	17.8	149	P. rapaceum	Hoarea	11
18-11-11	**Witzenberg Valley**	33	3	35.5	19	12	40.9	630	– unknown –		59
		33	4	10	19	12	45	640	P. trifoliolatum	Hoarea	21
21-11-11	**Elandsberg Reserve**	33	28	24.9	19	3	39.3	183	P. rapaceum	Hoarea	13
		33	28	19.6	19	3	35.3	170	P. longiflorum	Hoarea	18
		33	28	2.5	19	3	31.1	157	P. hispidum	Pelargonium	7
		33	27	58.1	19	3	26.7	152	P. rapaceum	Hoarea	10
		33	27	24.1	19	3	22.5	124	P. longiflorum	Hoarea	20
22-11-11	**Worcester**	33	24	37.3	19	12	42.9	292	P. rapaceum	Hoarea	27
		33	21	14.6	19	9	52.2	204	P. alchemilloides	Ciconium	33
		33	21	13.9	19	9	53.2	207	P. longifolium	Hoarea	18
		33	21	13.3	19	9	54.9	200	P. rapaceum	Hoarea	9
		33	36	37.2	19	26	54.3	381	P. trifidum	Jenkinsonia	7
		33	36	37.1	19	27	34.1	486	P. alternans	Otidia	16
23-11-11	**Bain's Kloof**								P. patulum	Glaucophyllum	14
									P. patulum	Glaucophyllum	45
		33	44	49.9	18	55	47.3	531	P. longicaule	Myrrhidium	6
		33	44	35	18	56	35	573	P. scabrum	Pelargonium	21
									P. triste	Polyactium	9
24-11-11	**Cape Town**	34	5	15.3	18	25	37.74	287	P. pinnatum	Hoarea	3
		33	55	5.22	18	23	59.23	295	P. rapaceum	Hoarea	6
		33	55	5	18	23	59.39	296	P. auritum	Hoarea	3
		33	56	44.61	18	23	47.3	331	P. rapaceum	Hoarea	11
25-11-11	**Jonkershoek**	34	0	11.1	18	59	37.8	526	P. longicaule	Myrrhidium	9
									P. patulum	Glaucophyllum	14
29-11-2011	**Newlands Ravine**	33	57	36	18	26	6		P. cucullatum	Pelargonium	10

Table 1.1. Fieldwork locations, results, and dates.

Herbaria

A total of 1348 specimen, belonging to 132 different species, were measured. Some species were only measured once or twice, whereas for others there were dozens of specimen available. For a full overview of all the measured species, see table 1.2.

Species	Section	Herbarium	Amount
P. abrotanifolium	Peristera	WAG	5
P. acetosum	Ciconium	BOL, NBG	7
P. aciculatum	Hoarea	STEU	6
P. acraeum	Ciconium	BOL, NBG	12
P. aestivale	Hoarea	BOL, STEU, NBG	10
P. album	Peristera	WAG	1
P. alchemilloides	Ciconium	WAG	7
P. alpinum	Pelargonium	WAG	1
P. anceps ssp. geniculatum	Jenkinsonia	STEU	7
P. appendiculatum	Ligularia	STEU, NBG	6
P. aridicola	Hoarea	BOL, STEU	8
P. aridum	Ciconium	BOL, WAG	6
P. aristatum	Hoarea	STEU	6
P. asarifolium	Hoarea	STEU	15
P. auritum var. auritum	Hoarea	BOL, STEU	18
P. auritum var. carneum	Hoarea	BOL, STEU	25
P. betulinum	Pelargonium	NBG, WAG	11
P. boranense	Myrrhidium	WAG	1
P. caledonicum	Hoarea	BOL, STEU	7
P. capillare	Campylia	WAG	1
P. carneum	Hoarea	STEU, NBG	19
P. carnosum	Otidia	WAG	2
P. caroli-henrici	Hoarea	STEU	5
P. caucalifolium ssp. caucalifolium	Myrrhidium	WAG	1
P. ceratophyllum	Otidia	BOL, NBG, WAG	6
P. citronellum	Pelargonium	WAG	1
P. confertum	Hoarea	BOL, STEU	8
P. connivens	Hoarea	STEU	3
P. cordifolium	Pelargonium	BOL	22
P. coronopifolium	Campylia	WAG	2
P. cortusifolium	Otidia	WAG	1
P. crithmifolium	Otidia	WAG	2
P. denticulatum	Pelargonium	BOL, WAG	12
P. dipetalum	Hoarea	BOL, STEU	19
P. dolomiticum	Jenkinsonia	WAG	2
P. echinatum	Otidia	BOL, WAG	23
P. elandsmontanum	Hoarea	STEU, NBG	3
P. elegans	Campylia	WAG	1
P. ellaphieae	Hoarea	BOL, STEU	14
P. elongatum	Ciconium	WAG	2
P. endlicheranium	Quercetorum	WAG	2
P. exstipulatum	Peristera	WAG	1
P. fasciculaceum	Hoarea	STEU	8
P. fergusoniae	Hoarea	BOL, STEU, NBG	16
P. fissifolium	Hoarea	STEU	31
P. flavidum	Hoarea	BOL, STEU, NBG	5
P. fruticosum	Pelargonium	BOL, WAG	21
P. fulgidum	Ligularia	NBG, WAG	5
P. fumariifolium	Hoarea	BOL, STEU, NBG	8
P. githagineum	Hoarea	BOL, STEU	3
P. glabriphyllum	Hoarea	STEU, NBG	2
P. glechomoidus	Peristera	WAG	6
P. glutinosum	Pelargonium	BOL	13
P. grandicalcaratum	Subsucculentia	WAG	1
P. graveolens	Pelargonium	WAG	5
P. grenvillae	Hoarea	BOL, STEU	7
P. grossularioides	Peristera	NBG	4
P. hermanniifolium	Pelargonium	BOL, WAG	9
P. hirtum	Hoarea	WAG	1
P. hispidum	Pelargonium	BOL, WAG	18
P. hypoleucum	Peristera	WAG	1
P. hystrix	Hoarea	BOL, NBG	9
P. incrassatum	Hoarea	BOL, STEU	30
P. inquinans	Ciconium	WAG	4
P. ionidiflorum	Peristera	WAG	1
P. karooicum	Quercetorum	WAG	1
P. ladysmithianum	Hoarea	STEU, NBG	4

(continued)

Species	Section	Herbarium	Amount
P. laevigatum ssp. laevigatum	Pelargonium	NBG, WAG	31
P. lanceolatum	Glaucophyllum	WAG	1
P. leipoldtii	Hoarea	STEU	20
P. leptum	Hoarea	BOL, STEU, NBG	5
P. longicaule var. longicaule	Myrrhidium	BOL, NBG	76
P. longiflorum	Hoarea	BOL, STEU, NBG	21
P. luridum	Magnistipulacea	WAG	18
P. luteum	Hoarea	STEU	6
P. magenteum	Otidia	BOL, WAG	22
P. minimum	Peristera	WAG	1
P. moniliforme	Hoarea	BOL, STEU, NBG	40
P. multibracteatum	Ciconium	WAG	34
P. myrrhifolium var. myrrhifolium	Myrrhidium	BOL	16
P. nanum	Nanum	WAG	3
P. nervifolium	Hoarea	BOL, STEU	14
P. oblongatum	Hoarea	BOL, STEU	8
P. odoratissimum	Peristerea	BOL, WAG	12
P. ovale	Campylia	WAG	5
P. pallidoflavum	Hoarea	STEU	8
P. paniculatum	Otidia	WAG	1
P. papilonaceum	Pelargonium	WAG	2
P. parvipetalum	Hoarea	BOL, STEU	6
P. patulum	Pelargonium	WAG	5
P. peltatum	Ciconium	WAG	4
P. petroselenifolium	Hoarea	BOL, STEU, NBG	7
P. pilosellifolium	Hoarea	STEU	21
P. pinnatum	Hoarea	BOL, STEU, NBG	50
P. praemorsum	Jenkinsonia	WAG	1
P. proliferum	Hoarea	STEU	23
P. pulchellum	Ligularia	WAG	3
P. punctatum	Hoarea	BOL, STEU	9
P. quarciticola	Hoarea	NBG	3
P. quercifolium	Pelargonium	BOL, WAG	23
P. quinquelobatum	Ciconium	WAG	17
P. radens	Pelargonium	WAG	1
P. radiatum	Hoarea	STEU	5
P. radulifolium	Polyactium	BOL	13
P. rapaceum	Hoarea	STEU	53
P. reflexipetalum	Hoarea	BOL, STEU	11
P. reflexum	Hoarea	STEU, NBG	6
P. reniforme	Peristera	WAG	4
P. scabrum	Pelargonium	WAG	9
P. senecioides	Jenkinsonia	WAG	9
P. setulosum	Pelargonium	WAG	1
P. spinosum	Subsucculentia	WAG	1
P. sublignosum	Pelargonium	BOL	10
P. suburbanum ssp. bipinnatifidum	Myrrhidium	BOL, NBG	10
P. suburbanum ssp. suburbanum	Myrrhidium	BOL, NBG	9
P. tenuicaule	Jenkinsonia	BOL, NBG	18
P. ternatum	Pelargonium	WAG	3
P. ternifolium	Hoarea	BOL, STEU	25
P. tetragonum	Chorisma	WAG	1
P. tomentosum	Pelargonium	BOL, WAG	8
P. tragacanthoides	Jenkinsonia	BOL	8
P. triandrum	Hoarea	BOL, STEU	9
P. tricolor	Campylia	WAG	1
P. triphyllum	Hoarea	BOL, STEU, NBG	16
P. triste	Polyactium	WAG	7
P. undulatum	Hoarea	BOL, STEU	13
P. vinaceum	Hoarea	STEU	13
P. violiflorum	Hoarea	BOL, STEU, NBG	8
P. vitifolium	Pelargonium	BOL, NBG	16
P. whytei	Myrrhidium	WAG	3
P. wuppertalense	Hoarea	STEU	12
P. zonale	Ciconium	WAG	5

Table 1.2. Names, numbers, locations and sections of all the specimen measured in the four different herbaria. Abbreviations: NBG = Compton Herbarium, WAG = Herbarium Vadense, STEU = Stellenbosch University Herbarium, BOL = Bolus Herbarium.

For accession numbers of all the analysed specimen, see Appendix A.

Nectar spur transformations

For both plants (*Pelargonium sidoides* and *Pelargonium zonale*) the flowers decreased both in weight and in spur length. A Student's t-test for paired samples was performed to check whether the decrease in weight or length was significant.

For both species, the decrease in weight was highly significant (*P. sidoides*: t-value = 14.82, degrees of freedom (df) = 19, p-value = $6.79 * 10^{-12}$; *P. zonale*: t-value = 8.85, df = 15, p-value = $2.43 * 10^{-7}$). On average, the *P. zonale*-flowers lost 90.46 percent of their weight, and the *P. sidoides*-flowers lost 84.26 percent. The average loss of weight therefore was 87.36 percent.

The decrease in nectar spur length was less pronounced, but still significant (*P. sidoides*: t-value = 5.16, df = 19, p-value = $5.54 * 10^{-05}$; *P. zonale*: t-value = 2.52, df = 15, p-value = $2.38 * 10^{-2}$). The average decrease in nectar spur length for *P. zonale*-flowers was 4.12, while *P. sidoides*-flowers lost 6.07 percent of their spur length. The average loss of spur length was 5.1 percent.

However, results varied per species and level of pressure. When comparing nectar spur lengths of untransformed and transformed *P. zonale* flowers for each of the levels of pressure used (heavy pressure, normal pressure, light pressure, no pressure), none of the decreases in spur length were significant ($p > 0.05$). A significant change in nectar spur length was only detected when grouping all measurements of *P. zonale* together. In contrast, the decrease in spur length of *P. sidoides* flowers was significant when using normal pressure, light pressure, and no pressure, as well as when all measurements were grouped.

To account for this loss in spur length, all nectar spur lengths obtained from herbaria specimen were multiplied with 1.051 before any successive analyses were performed.

Nectar spur length dataset

All nectar spur lengths, measured both in the field and in the four herbaria, were combined in one dataset. This dataset was complemented with spur length measurements from the scientific literature, mostly from the books of Van der Walt (1977), Van der Walt & Vorster (1981, 1988) and Marais (1981). This resulted in one dataset with spur length measurements for 186 *Pelargonium* species, subspecies, and varieties. This dataset was used in all following statistical and phylogenetic analyses.

The main component of the dataset is the median nectar spur length per species (see Appendix B). However, as Hardy & Linder (2005) noted, the mean value of a continuously varying trait does not necessarily convey more information about the ecology of a species than any other value within the range of trait variation of the species. Therefore, for all species, other character values were calculated as well: the minimum value, the maximum value, the median, and the range of character values. If enough spur lengths were measured, the standard deviation and the variance were calculated as well. As the mean value of a continuous trait may be heavily affected by large outliers, and median values do not suffer as much from this problem, in this thesis the main measure of nectar spur length is its median, rather than its mean.

Statistics

To gain further insight in nectar spur lengths distributions and characteristics, the available dataset, consisting of measurements from wild populations, herbaria, and the scientific literature, was analysed using R.

Spur length distributions

The average median nectar spur length is 2.03 centimetres. Average minimum, maximum, and range values are 1.42, 2.79, and 1.38 centimetres respectively. The distribution of these four standard spur measurements can be seen in figure 1.5.

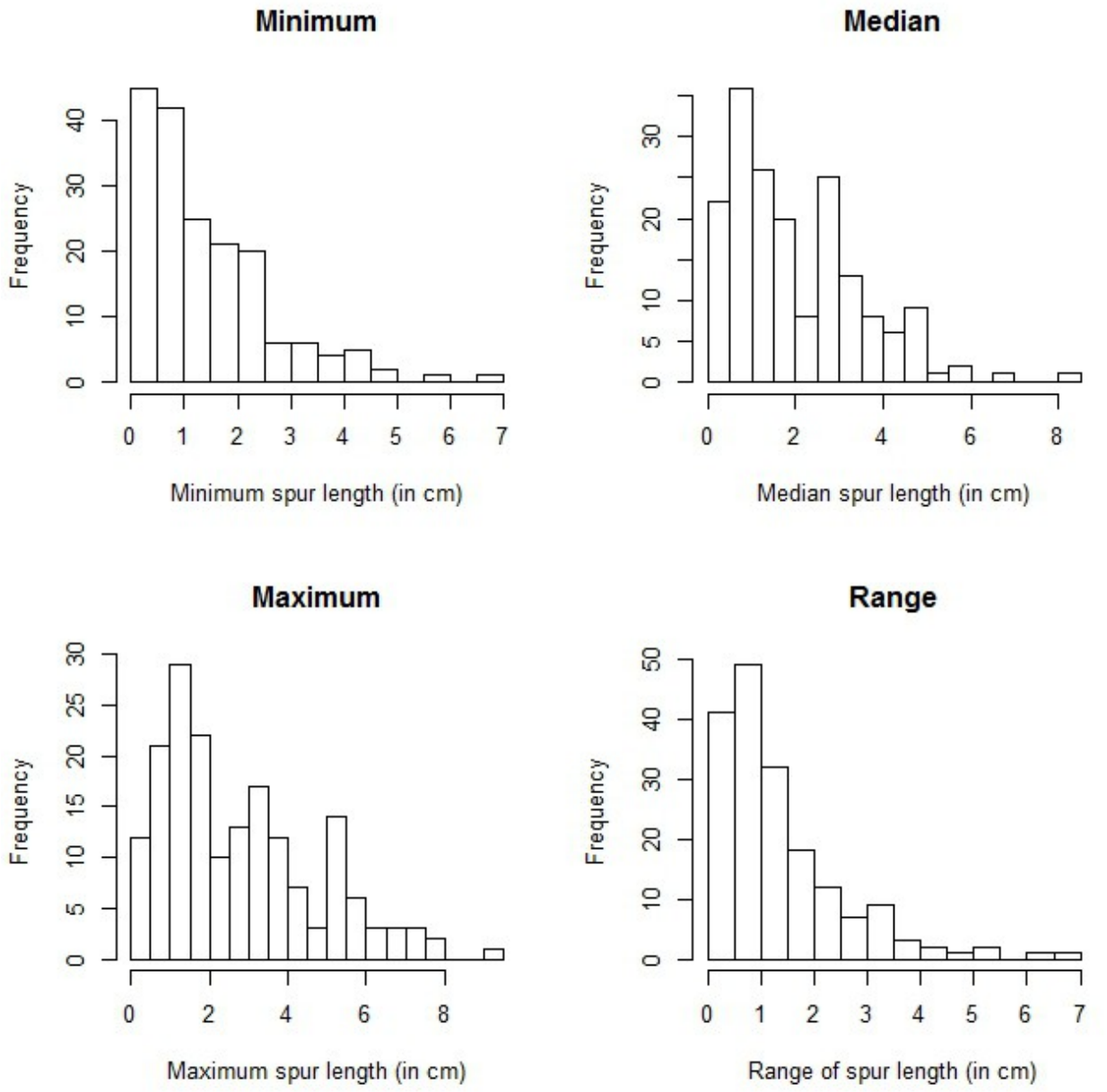

Figure 1.5. Histograms of the minimum, median, maximum recorded nectar spur length and the range of nectar spur length across all species.

All four distributions seem to follow the same pattern: low measures are rather common, whereas larger values are much rarer.
When looking at individual species, distributions of nectar spur lengths approach normality. The only exception is *P. myrrhifolium*; its nectar spur distribution seems to show two distinct peaks. See figure 1.6 for histograms of spur lengths of the four most sampled species in the field.

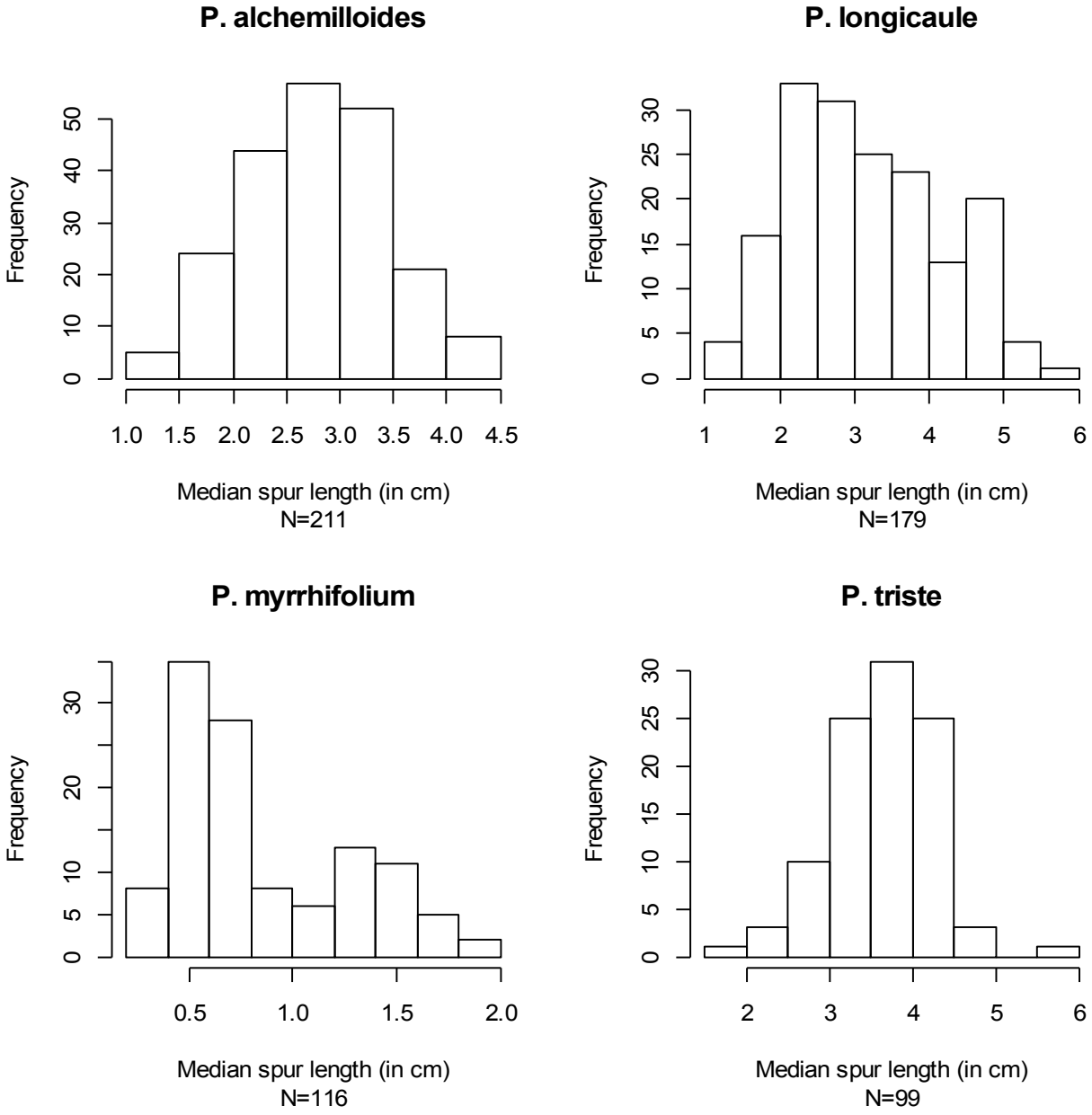

Figure 1.6. Distribution of spur lengths of four most sampled species in the field. For graphical purposes, the x-axis, depicting median spur length in centimetres, differs per species.

Interpopulational variation

A normal distribution is approached even further when only the distribution of nectar spur lengths within a population is considered. See figure 1.7 for histograms of spur length within the four largest populations of the most-sampled species (*Pelargonium alchemilloides*).

Different species have different levels of inter-population variation. Of the four most-sampled species, spur lengths are quite similar between different populations of *P. alchemilloides* and *P. triste*. However, larger differences exist between the different populations of *P. longicaule* and *P. myrrhifolium* (see figure 1.8).

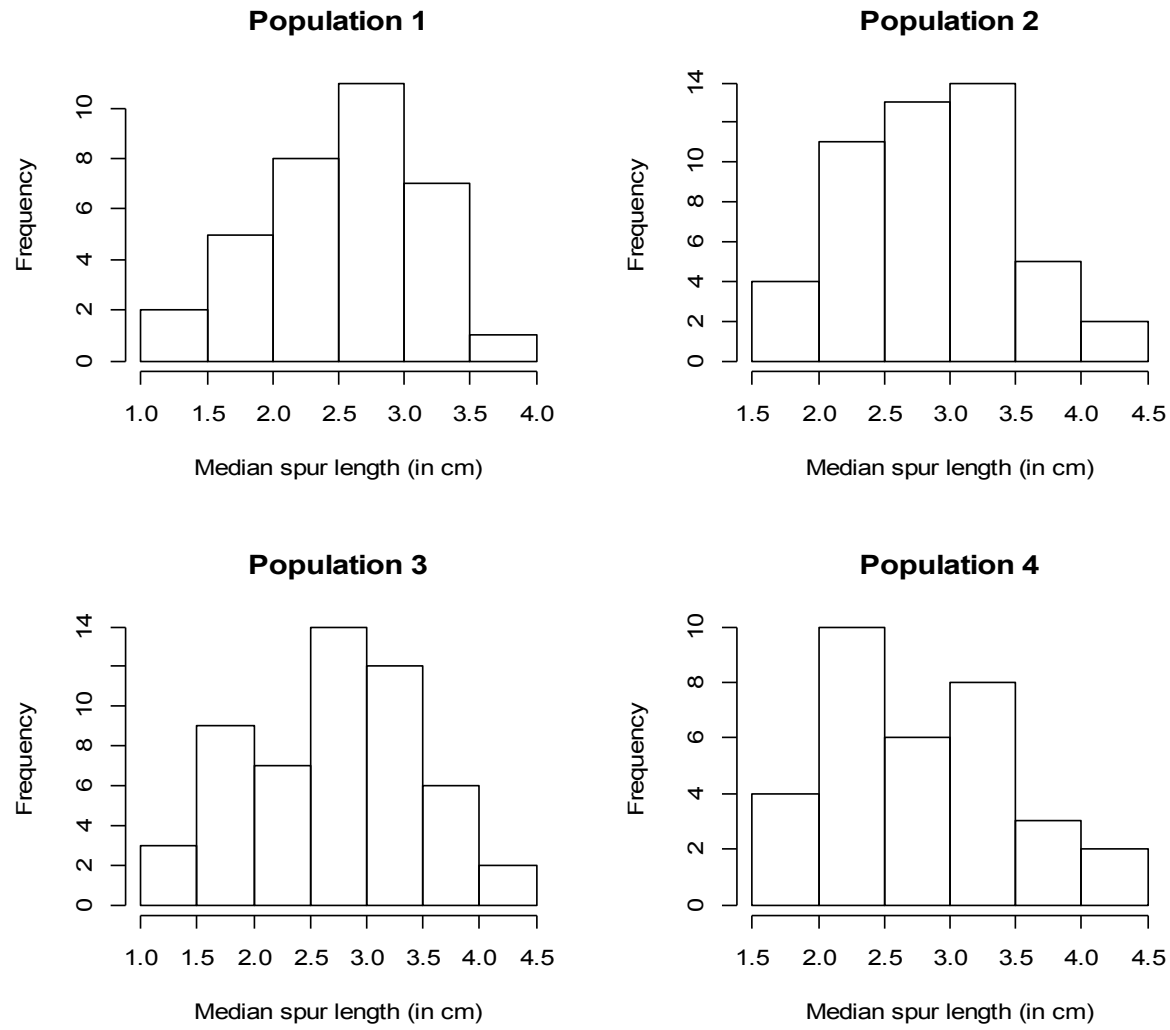

Figure 1.7. Distribution of nectar spur length of four largest populations of *P. alchemilloides*.

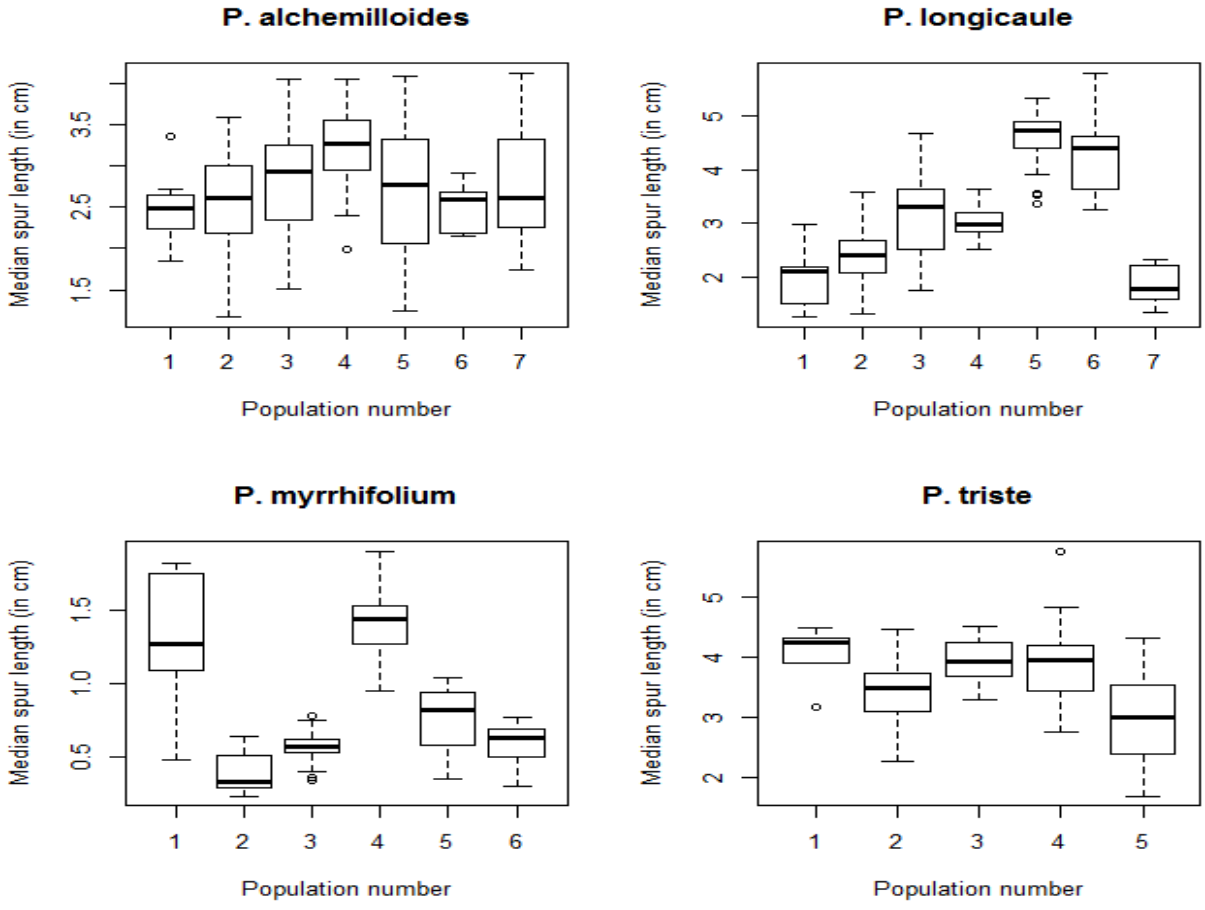

Figure 1.8. Differences in spur lengths between populations of most-sampled species. The small circles depict outliers in the spur length range.

The amount of variation of nectar spur length within populations differs greatly per species and population. To make this visible, standard deviations of nectar spur length per population were calculated and plotted in a histogram (see figure 1.9). As can be seen, for the vast majority of species, the standard deviation - and therefore the level of variation within a population - is quite low (<0.1 centimetres). However, for some species the standard deviation is as high as 0.6 centimetres.

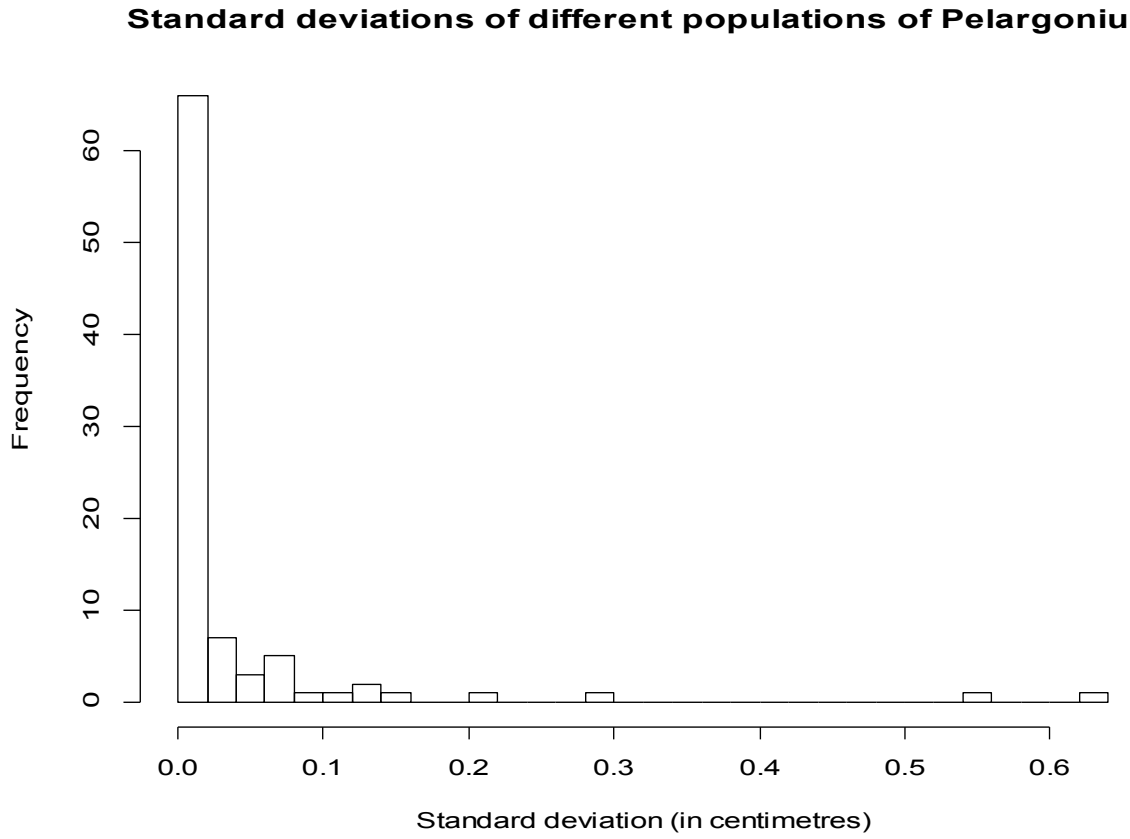

Figure 1.9. Standard deviation of nectar spur length for each measured *Pelargonium* population. A total of 90 populations were analysed.

Variation within individuals

As mentioned previously in this chapter, to assess the variation within individuals, 10 nectar spurs per individual were measured for 10 *P. cucullatum*-plants, all belonging to the same population. The results are depicted in figure 1.10. It can be seen that there is a lot of variation between and within individuals, even though all these individuals were all fully grown, and living only a couple of metres from each other.

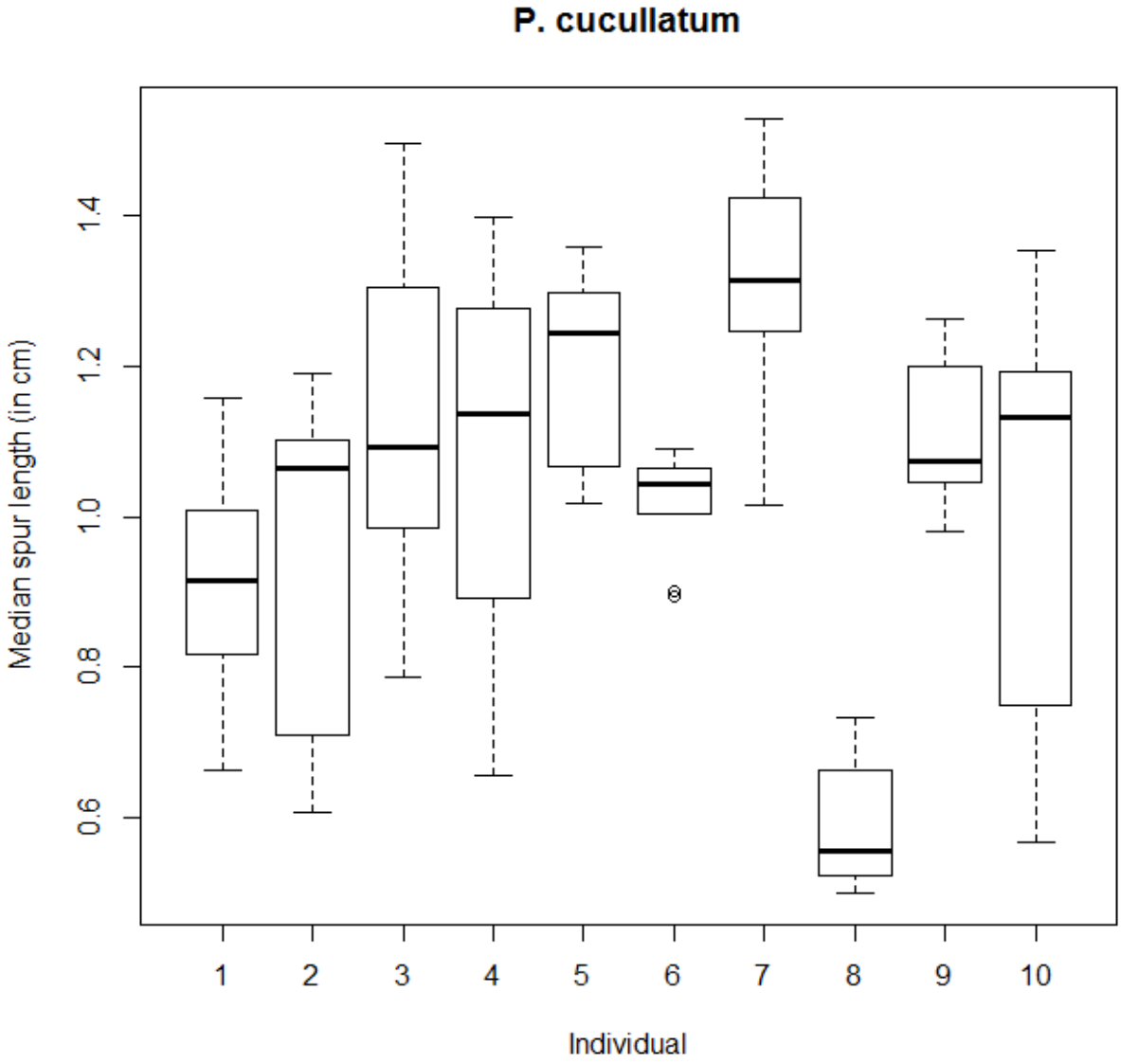

Figure 1.10. Differences in spur lengths within and between 10 *P. cucullatum*-plants all belonging to the same population.

Correlation

To test for the correlation between the different measures of spur length, Pearson's Product-Moment Correlation Coefficient was calculated. This coefficient may vary from 1 (implying perfect correlation between two parameters) to -1 (indicating perfect but reciprocal correlation), where a value of 0 indicates statistical independence (Moore & McCabe 2009).
Correlation coefficients were calculated for combinations of the four different measures - minimum, median, maximum and range of spur length - and the four different groups - field measurements grouped per population, field measurements grouped per species, herbarium measurements, and measurements from the combined dataset. This resulted in a total of 24 different correlation coefficients. All these coefficients were highly significantly different from 0 (p-value << 0.0001), with the exception of the correlation coefficient of the minimum and range value for field measurements grouped per species (p-value < 0.01), the herbarium sample (p-value < 0.01), and the combined dataset (p-value < 0.001). As these correlation coefficients are still sufficiently significantly different from 0, all coefficients can be analysed.

The correlation coefficients per measurement and per group can be seen in table 1.3 and figure 1.11.

Field – per population		Field – per species		Herbarium		Combined	
Median-range	0.7949163	**Median-range**	0.7375303	**Median-range**	0.5010715	**Median-range**	0.5365247
Max-min	0.9184941	**Max-min**	0.7676157	**Max-min**	0.7789528	**Max-min**	0.8070957
Median-min	0.9563757	**Median-min**	0.9061291	**Median-min**	0.9326783	**Median-min**	0.9364
Median-max	0.9802983	**Median-max**	0.9432993	**Median-max**	0.8915225	**Median-max**	0.9218366
Range-min	0.6197957	**Range-min**	0.4773631	**Range-min**	0.2652092	**Range-min**	0.2852703
Range-max	0.8795808	**Range-max**	0.9295718	**Range-max**	0.8112127	**Range-max**	0.7956334

Figure 1.3. Pearson's correlation coefficient for different measures of nectar spur length and different groups. All values are highly statistically significant (p-values <0.01).

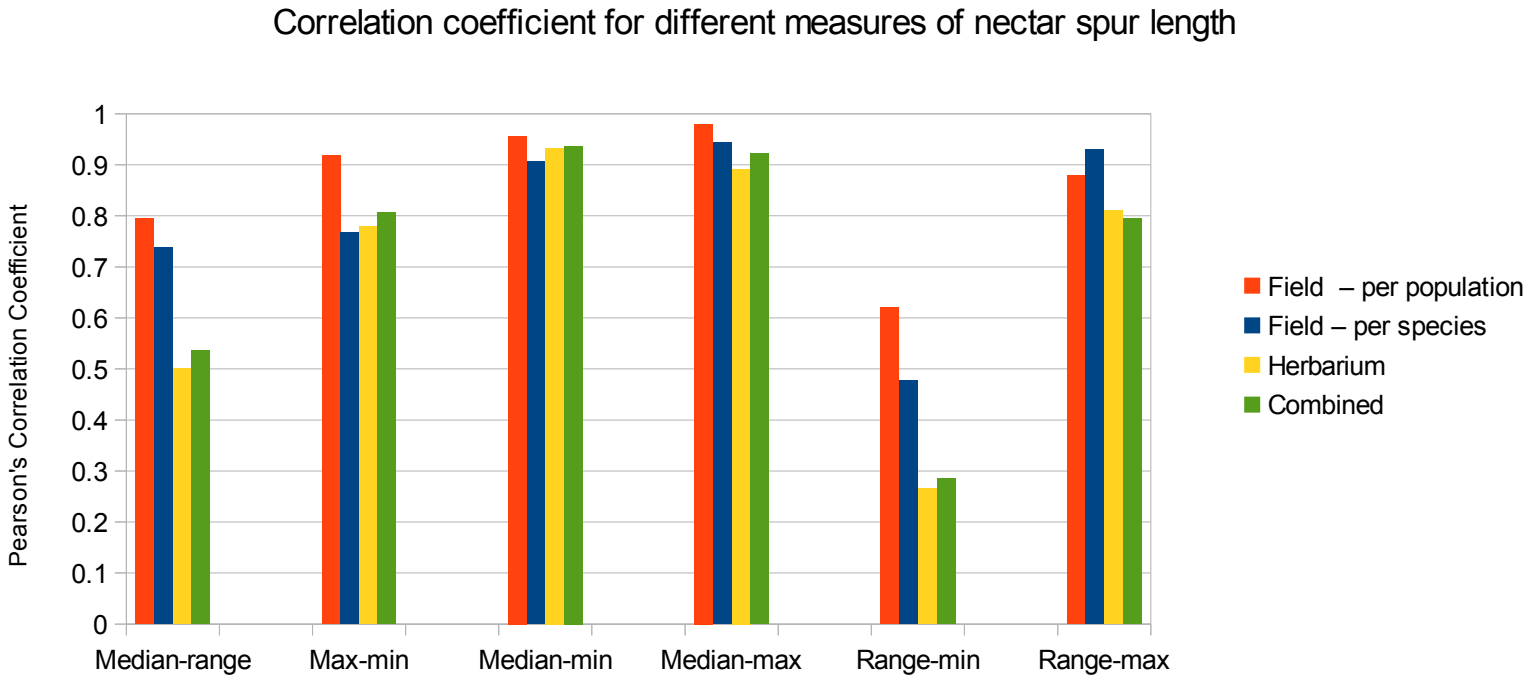

Figure 1.11. Pearson's correlation coefficient for different measures of nectar spur length and different groups.

Several conclusions can be drawn based on these correlation coefficients.
First of all, there is a moderate to strong correlation between the median and the range of nectar spur length of a certain species. The correlation coefficient is approximately 0.5 for the herbarium and combined dataset, indicating a moderate correlation, and 0.7-0.8 for the field measurements, implying strong correlation. This means that when a certain species has a relatively high nectar spur length, there is a high probability it also has a relatively large nectar spur range. To put this in more simple terms, a species that has large nectar spurs probably also has much intraspecific variation in nectar spur length, whereas a species with smaller nectar spurs has less variation.
A second observation that can be made is that there is a moderately strong to strong correlation between minimum and maximum spur length, as well as between minimum and maximum spur length on the one hand and range and median spur length on the other hand. Correlation coefficients range between 0.78 and 0.98, with many of them exceeding 0.90 (see figure 1.12). This indicates

strong correlation between the different measures of spur length; larger maximum nectar spurs indicate larger minimum, range, and median spur lengths, and vice versa.
However, a third conclusion is that there is one interesting exception to the aforementioned point: the correlation between minimum spur length and the range of spur length is rather weak for all four groups. The highest coefficient for this correlation is 0.62, and the lowest is as low as 0.27 (see figure 1.12). Based on minimum nectar spur length, one could therefore conclude that larger nectar spurs do not necessarily increase the variation in nectar spur length. However, as can be seen from the correlation coefficients of the median and maximum values with the range of spur length, larger nectar spurs are in fact quite strongly correlated with larger variation. This indicates that the minimum nectar spur length of a species may not be a suitable measure of the spur length variation of this species.
A final conclusion that can be drawn is that for all different measures of nectar spur length, the correlation coefficients for the spurs measured in the field are (much) higher than the coefficients for spurs belonging to the herbarium sample or the combined dataset. Sometimes the differences are small (0.02), for other measures of spur length the differences are quite large (up to 0.35). This may be due to the fact that for most species, field measurements are based on dozens of individuals, whereas measurements in the herbaria and in the scientific literature sometimes are only based on a few specimen. In a similar way, for nectar spurs collected in the field, spurs grouped per population have slightly larger correlation coefficients than spurs grouped per species.

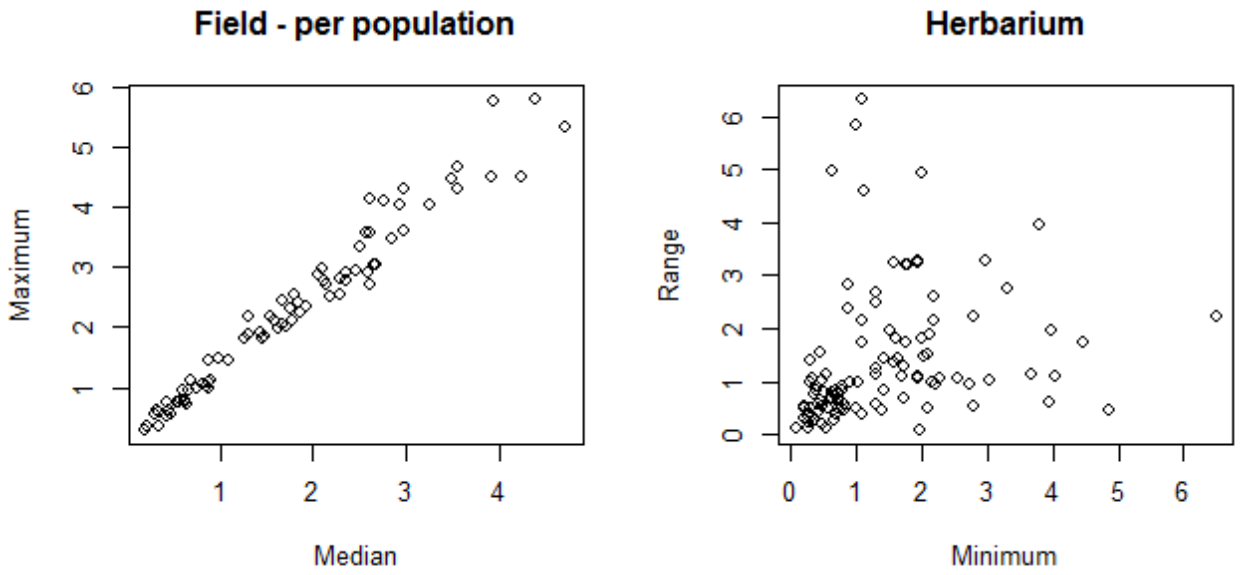

Figure 1.12. Plots of the strongest (left) and weakest (right) correlation found between different measures of nectar spur length per species. The left figure shows the correlation between maximum and median spur length as measured for individual populations in the field and has a correlation coefficient of 0.98. The right figure shows the correlation between minimum spur length and the range of spur length as measured from herbarium specimen, with a correlation coefficient of 0.27. The axes are measured in centimetres.

Discussion

Spur length variation within species, populations, and individuals

The first hypothesis of this thesis states: '*Pelargonium* nectar spur length is constant at the population level.'. After measuring and analysing hundreds of nectar spurs of dozens of species, both in the field and in herbaria, I am forced to partially reject this hypothesis; there are considerable amounts of intraspecific and intrapopulational variation in nectar spur length in *Pelargonium*. However, the amount of variation varies widely between different species and even between different populations.

The range of spur length (the difference between the minimum and the maximum spur length of a certain species) shows much variation for different species. For the majority of the species it is around 2 centimetres (see figure 1.5), but for some species, the range may be as high as 6 centimetres. Furthermore, spur length does not only vary within complete species, but also within and between population of a certain species. For long-spurred species such as *P. longicaule* and *P. triste*, the difference between the longest and the shortest nectar spur within a population may be up to three to four centimetres (see figure 1.6). And for the short-spurred *P. myrrhifolium*, spur length variation within one population was almost 2 centimetres - considering that the median spur length of this species is 0.7 centimetres, this is quite a large range of variation.
These levels of variation can not simply be explained by the presence of a few outliers. The standard deviation for some species and populations is remarkably high: for *P. longicaule* as a complete species, it is as high as 1.1 centimetres, while for certain populations of *P. alchemilloides* and *P. triste* it is in the range of 0.2 to 0.6 centimetres. However, it is important to realize that for most species, the standard deviation of spur length was in fact quite low (< 0.1 centimetres, see figure 1.9). This indicates that not only variation in nectar spur length exists between species, but that the level of variation is highly variable as well, for different species.

In fact, even within an individual, considerable spur length variation may exist (see figure 1.10). For one population of *P. cucullatum*, 10 flowers per individual for 10 individuals were analysed. This showed that the amount of spur length variation within an individual may vary widely; for some individuals, the range in lengths was as low as 0.1 centimetres, whereas for other individuals this range approached 1 centimetre. Even though the 10 individuals all belonged to the same population, and were growing literally within 5 metres from each other, there was much variation in spur length between the different plants.

As nectar spur length and pollinator proboscis length are thought to need to be closely matched (e.g., Johnson & Steiner 1997, Whittall & Hodges 2007, Givnish 2010) in order for successful pollination to occur, these levels of variation seem counter-intuitive. Five hypotheses may explain the relatively large variation in nectar spur lengths in some species.

1) The Geographic Mosaic Theory of Coevolution (Pauw *et al*. 2008). This theory postulates that different populations of a certain plant species are under different selective pressures regarding the length of their nectar spurs. Some plants may be locked in an 'arms-race' with their pollinators, leading to coevolution and steadily increasing nectar spurs and proboscises lengths (Ennos 2008, Rodríguez-Gironés & Llandres 2008, Johnson & Anderson 2010). In other species, spur lengths are heavily influenced by plants adapting to a certain pollinator: the pollinator shift model (Whittall & Hodges 2007), causing nectar spurs to rapidly change in length. As different processes occur in different populations and different species, this leads to a patchwork of populations. Pauw *et al.* state that some of the patterns predicted by such a patchwork include that "populations will differ in the traits shaped by an interaction" and that "high levels of polymorphism will be maintained in

some populations". In fact, one of the conclusions from the study of Pauw and colleagues is that remarkable levels of nectar spur/tube length variation exists in a few South African plant species, and one of these species is *P. longicaule.* This Geographic Mosaic Theory of Coevolution may explain the levels of spur length variation within and between populations found in this thesis.

2) Different pollination strategies exist per individual plant. Pollination is usually seen as a process that acts on the population level (Struck 1997, Johnson 2010): most individuals within a population are pollinated by one (or a few) pollinators. Differences between populations may exist, but within a population, there tends to be one primary pollinator. The levels of nectar spur length variation of some populations of *Pelargonium* could suggest something else; within populations, different species may adopt different pollination strategies, and therefore attract different pollinators. It is possible that the individuals with the shortest spurs in a population attract generalist pollinators, such as bees (Struck 1997), whereas species with longer spurs attract more specialist pollinators, such as hawkmoths and long-proboscid hovering flies (Goldblatt & Manning 2000, Borrell 2005). This may be a first step to reproductive isolation and therefore to speciation (Schluter 2001). However, it may also be possible that several pollination strategies keep existing within a population. In that case, the success of a certain strategy will be based on the presence and preference of pollinator species, which may differ within and between years (Johnson 2010). An indication of the multiple pollinators hypothesis could be that more than one peak exists in the nectar spur length distribution; as the spur length distributions of all species are characterised by only one peak, therefore this hypothesis may have to be rejected. The only exception is the spur length distribution of *P. myrrhifolium* (see figure 1.6), which seems to show hints of a second peak.

3) Nectar spur length and proboscis length are not exactly matched. It is possible that small levels of spur length variation do not pose an obstacle to visiting pollinators. This hypothesis is congruent with observations from a large number of studies where spur or tube length and pollinator proboscis or tongue length do not match exactly: in the orchid *Satyrium* (Johnson *et al.* 2011), in the long-tubed iris *Lapeirousia anceps* (Pauw *et al.* 2008), in the *Disa draconis* complex (Johnson & Steiner 1997), and in *Gladiolus* (Goldblatt & Manning 1999). All these studies are focussed on the Cape Floristic Region, and worldwide there are numerous other examples. In fact, Anderson *et al.* (2010) showed that trait mismatches may be the rule, rather than the exception, of plant insect interactions. This may indicate that the relatively low levels of intrapopulational spur length variation typical for most *Pelargonium* species do not reflect a similar variation in pollinators; as long as the variation is low, one pollinator may visit all the individuals within a population without problems.

4) Similar variation exists in proboscis length. Just as in the nectar spurs, a level of variation may exist in pollinator proboscis or tongue length. From several studies it is in fact known that considerable variation in proboscis length exists within species of pollinators (e.g., Pauw *et al.* 2008, Combs & Pauw 2009, Johnson *et al*. 2011). According to this hypothesis, individual pollinators with slightly longer proboscises may focus on the plants in a population with slightly longer nectar spurs, whereas pollinators of the same species with shorter proboscises focus on the shorter-spurred individuals in the same population. Such reciprocal variation in spur and proboscis length may be the first step to speciation (Rodríguez-Gironés & Llandres 2008, Rodríguez-Gironés & Santamaría 2010).

5) There is no correlation between nectar spur length and proboscis length in (some) *Pelargonium* species. One of the central premises of this thesis is that the spur length of a *Pelargonium* species is a proxy for its respective pollinator. This premise is corroborated by the vast majority of nectar spur studies in the scientific literature, both general studies (Hodges & Arnold 1995, Hodges 1997) as well as studies on specific genera and species (e.g., Borrell 2005, Whittall & Hodges 2007). In fact, several studies on certain *Pelargonium* species also state that the length of a nectar spur is (loosely) correlated with the length of the proboscis of its respective pollinator (Struck 1997, De Wet *et al*. 2008, Combs & Pauw 2009, Van der Niet & Bakker (unpublished)).

However, based on the levels of spur length variation found in some species and populations in this thesis, the possibility that spur length and proboscis length in *Pelargonium* are not as closely matched as in other genera can not completely be ignored. It is absolutely not unthinkable that other floral characteristics, such as for instance flower colour, nectar components or length and position of stamens (Marais 1999), are more important than nectar spur length in attracting the attention of a pollinator. This would allow other forces to have a larger evolutionary effect on nectar spur length than pollinator pressure. For, as Pauw and colleagues (2008) state "Indeed, the majority of selective forces that act on proboscis and tube [= nectar spur] length remain unknown". It may even be possible that the range of spur length variation reflects phenotypic rather than genetic differences, as it is known that spur lengths tend to be influenced by environmental factors such as sunlight (E.M. Marais, pers. comm.).

Unfortunately, as there are hardly any direct pollinator observations for *Pelargonium*, these explanations are mostly based on speculation. Much fieldwork will need to be performed before distinguishing between these hypotheses is possible (see also chapter 3).

Different measures of spur length and their correlation

In this thesis different measures of nectar spur length were calculated and analysed, in order to gain insight in the distribution and variation of spur lengths in natural populations. For most statistical and phylogenetic analyses in this thesis, four of these measures were used: minimum, median, maximum and the range of spur length. Pearson's correlation coefficient was calculated for combinations of all of these four measurements. Three important conclusions can be drawn based on the results.

First of all, there is a strong, positive and significant correlation between median and maximum spur length on the one hand and the range of spur length on the other hand. The correlation coefficient generally lies between 0.75 and 0.90, implying strong correlation between the two measures of spur length. This indicates that when a species has relatively long nectar spurs, there tends to be more intraspecific variation than when a species has shorter spurs. Examples of this are easy to find: *P. longicaule* has fairly long nectar spurs and much larger levels of variation than the shorter-spurred *P. senecioides*.

This correlation between median/maximum spur length and the range of spur length is found when comparing whole species, but also when comparing individual populations of *Pelargonium* species. In fact, the correlation coefficients are slightly higher (approximately 0.04) when the individuals are grouped per population rather than per species (see figure 1.11). Considering the pollinator fidelity hypothesis, these results are somewhat remarkable. This hypothesis states that the function of nectar spurs is to exclude certain pollinators and to attract others (Kay & Sargent 2009). As pollinators with longer proboscises tend to be specialist pollinators (Borrell 2005), whereas shorter-spurred pollinators such as bees are more generalists (Struck 1997), elongating nectar spurs will increase the chance of attracting specialist pollinators, and therefore the chance that the pollen of a specific individual will be transmitted to individuals of the same species (Rabosky & McCune 2009, Johnson & Anderson 2010). In other words: longer nectar spurs increase the pollinator fidelity. However, if longer nectar spurs do evolve to increase the pollinator fidelity, one would expect longer spurs to show lower levels of variation, rather than higher. In order to accept or reject this hypothesis, more pollinator observations are necessary.

A second conclusion that can be drawn based on the correlation analyses is that the correlation between minimum spur length and the range of spur length is much lower than the correlation

between any other combination of spur measurements; it varies between 0.25 and 0.60. Even though this correlation coefficient is still positive (and significant), it indicates a completely different pattern than the correlation between median and range. Based on the correlation between minimum and range, one could conclude that longer nectar spurs do not necessarily increase the amount of spur length variation. However, as can be seen from the correlation coefficients between median/maximum and range, there is in fact quite a strong correlation between spur length and variation in spur length. Based on this, it can be concluded that minimum nectar spur lengths are not an ideal indicator of the total existing variation. This could be a reason to exclude minimum spur lengths from phylogenetic analyses.

It is interesting to think about the biological aspects of this observation. As minimum spur lengths are not a good indicator of existing variation within a species or population, and therefore minimum spur lengths are more constant across populations and species than other measures of spur length, does this mean that individuals with the shortest spurs in most *Pelargonium* species can in fact be pollinated by the same generalist pollinators? Maybe minimum spur lengths are much more conserved than other measures of spur length across the phylogenetic tree of Pelargonium, and most of the nectar spur evolution occurs in the longest spurs. This question will be further addressed in the following chapter.

The third conclusion is that correlation coefficients for all combinations of spur measurements tend to be higher for spurs measured in the field, and especially for the field measurements grouped per population (rather than per species). Patterns and correlations that are clearly visible in field populations remain hidden when analysing herbarium specimen or records from the literature, with differences in correlation coefficients sometimes as high as 0.4. This is clear evidence for the theory that selection acts on nectar spur lengths on the population level.

Implications for future research

Based on these results, three recommendations for future nectar spur research, either on *Pelargonium* or on other genera, can be formulated.

1) Measure a considerable amount of individuals per species. In many studies, only several individuals are measured per species. For example, Whittal & Hodges measured nectar spurs of 10-52 individuals and 1-5 population per *Aquilegia* species and Combs & Pauw analysed nectar spurs of 6 individual *Pelargonium* plants. The results of this thesis suggest that this way large amounts of variation may be missed, especially when analysing long-spurred species.

2) Measure nectar spurs in natural populations. Pollination is a population process, and therefore nectar spurs should be measured and analysed in the context of their respective populations. If this is not possible, the spurs should at least be measured in the field. By analysing spur length data obtained from herbaria or the scientific literature, much information will be lost.

3) Use several measures of spur length. Nectar spurs are complex structures, which vary widely across populations. They can not simply be caught in one number, such as the mean; instead, it is best to use different measures of spur length, to gain full insight in the existing distributions and variation. Or, as Hardy & Linder (2005, in Hardy 2006) put it: "the average eco-parameter value for a species does not necessarily encompass more information about the historical ecology of the species than any other value within the variation range of the species".

Chapter 2:
Nectar spur evolution

Introduction

The first chapter of this thesis was mainly focussed on the distribution and statistics of nectar spur lengths of *Pelargonium* species in natural populations. The aim of this chapter is to gain more insight in the evolutionary processes that shaped these spur length distributions. Special focus is given to how nectar spur length has changed during the evolution of *Pelargonium*, and if and how nectar spur evolution influenced speciation rates. This is done by testing the second and the third hypothesis of this thesis.

The second hypothesis states: 'Pollinator-switches and clade-proliferation are linked in clades A2HLM and A2POC'. Clade A2POC contains the sections *Polyactium*, *Otidia*, and *Cortusina*, and clade A2HLM consists of the sections *Hoarea*, *Ligularia*, and *Magnistipulaceae*. Section *Hoarea* is by far the largest section of *Pelargonium*, containing more than 80 species (Touloumenidou *et al.* 2003), as well as large levels of floral nectar spur variation. Clade A2POC is much smaller, containing only 25 species. According to the study of *Pelargonium* pollination by Bakker and colleagues (2005), a switch from bee to long-proboscid hovering flies occurred in the A2HLM-clade (A2b in Bakker *et al.* 2005), whereas no such switch occurred in clade A2POC. This could indicate a correlation between pollinator-switch and clade-proliferation. However, a similar situation exists in clade C2 (switch from bees to long-proboscid hovering flies) and C1 (pollination by bees), but these two clades are approximately the same size. This makes a correlation between pollinator-switches and clade-proliferation less likely.

In the scientific literature, pollinator-switches are considered to play a highly important role for clade-proliferation (Kay & Sargent 2009, Johnson 2010). During the development of their model, Grant (1949) and Stebbins (1970) placed much emphasis on the importance of pollinator-switches. Many studies have confirmed their hypothesis by demonstrating that switches in pollinator use have occurred frequently during clade radiation, both globally (Givnish & Sytsma 1997, Weller & Sakai 1999) and in South Africa (in almost every speciation event in *Disa*, a pollinator switch took place (Johnson *et al.* 1998), and in *Gladiolus* and *Babiana* (Iridaceae), Goldblatt & Manning (2006) estimated at least one switch for every five to six species (Johnson 2010).

Based on this, it seems clear that pollinator-switches are highly important for clade-proliferation, but evidence does also suggest that both are not correlated per se in *Pelargonium*. In this chapter, I will try to test this hypothesis by analysing nectar spur evolution with several phylogenetic methods contained in the R statistical language (R Core Development Team 2012). It is important to note that, as I do not have any pollinator data, I will look for a correlation between nectar spur evolution and clade-proliferation, rather than a correlation between pollinator-switches and clade-proliferation. See chapter 3 for more information about the relationship between nectar spurs and pollinators.

The third hypothesis considers a direction in spur evolution: 'In the evolution of nectar spurs in *Pelargonium*, there was a trend towards longer spurs'. According to Bakker *et al.* (2005), switches from pollination by bees to pollination by long-proboscid hovering flies, moths or even birds occurred several times. All these switches involve elongation of the nectar spur. Based on this, it is tempting to say an evolutionary trend exists towards longer spurs, similar to the trend in *Aquilegia* (Whittal & Hodges 2007).

There may be a rational explanation of such a trend towards longer nectar spurs: longer spurs could induce pollinator fidelity (Kay & Sargent 2009, see also Chapter 1). As a direct result of the long nectar spurs of some *Pelargonium* species, insects with short proboscids are unable to drink nectar from the tube, therefore they will not act as pollinators for these long-spurred species. Insects with long proboscises, however, are still able to reach the nectar in the spur, and will still serve as pollinator to these species. Therefore, by elongation of the nectar spur, plants are able to exclude pollinators with short proboscises, which are often generalist pollinators (Struck 1997) and will only be visited by specialist pollinators with long proboscises. The process of increasing pollinator fidelity by increasing floral specialization has been demonstrated in several studies, both theoretically and empirically (Sargent 2004, Stang 2007, Rabosky & McCune 2009, Kay & Sargent 2009).
Increasing pollinator fidelity may be especially important for species that cannot afford the waste of their pollen: species that live in harsh conditions with little resources and in small populations. This describes most species of the section *Hoarea* (Van der Walt 1977, Van der Walt & Vorster 1981, 1988) in clade A2HLM. Therefore it may not be surprising that most of the nectar spur evolution can in fact be found in this clade.

However, several factors suggest the proposal of the existence of such a trend should be handled with caution. First of all, the analysis of Bakker and colleagues (2005) was performed using the rather simple reconstruction method of parsimony over one single tree. In this thesis, I incorporated phylogenetic and model uncertainty by using several methods to optimise spur lengths over a whole range of trees, which may lead to significantly different answers.
Furthermore, in several terminal branches reversals in spur length evolution occurred in *Pelargonium*. The most remarkable example is *P. hirtum*, with a spur length of only 1-5 millimetres, whereas its sister species, *P. appendiculatum*, has the longest spurs in the whole genus, stretching to more than 75 millimetres. Additionally, in four different species nectar spurs disappeared altogether. However, this mostly happened to plants that have dispersed from South Africa, and therefore are subject to completely different pollination pressures than the plants in the Cape. It also seems worth noting that these reversals only occurred in distal nodes of the tree, whereas shifts to longer spurs occurred on more basal nodes as well (Bakker *et al.* 2005).
Finally, one should have quite a good reason to propose the existence of a trend towards longer or bigger structures in general (Gould 1997). For, as Gould so elegantly put it, “Our strong and biased predilection for focusing on extremes (and misconstruing their trends as surrogates for a totality), rather than documenting full ranges of variation, generates all manner of deep and stubborn errors.”

To summarise, based on the available information, it is quite difficult to conclude whether or not there is an evolutionary trend towards longer spurs in *Pelargonium*. Once again, I will test this hypothesis by using phylogenetic methods incorporated in R. As these methods need a phylogenetic tree, I will first describe how such a tree was developed, and what the resulting phylogenetic tree looks like. Afterwards, the different phylogenetic methods and their results are discussed, followed by a discussion about the hypotheses and premises of this thesis.

For R scripts of the most important analyses performed in this chapter, see Appendix C.

Phylogenetic inference

The main focus of this project is the collection of the nectar spur measurements in the field and the phylogenetic analysis of nectar spur evolution. For this second step a phylogenetic tree of *Pelargonium* is needed. As the process of creating this tree is not the main focus of this project, I will only briefly describe the methods used for phylogenetic inference.

Dataset

The dataset used for the creation of the phylogenetic tree consisted of 232 taxa. Several of these taxa consisted of subspecies, varieties, *Pelargonium* species that have not yet been described or species that are represented more than once in the dataset. With the removal of these taxa, DNA data was available for 186 unique *Pelargonium* species.

The dataset consisted of 4 partitions with in total 1846 characters. These four partitions are:

1) cpDNA *trnL-F* (1029 sites),
2) nrDNA *ITS* (613 sites),
3) mtDNA *nad1* (168 sites), and
4) 30 plastid indels.

For a complete description of the data used, please see Bakker *et al.* (2004), Jones *et al.* (2009), Van Proosdij *et al.* (in review), and Ringelberg *et al.* (in preparation).

Analysis

The dataset was analysed using MrBayes version 3.2.1 (Huelsenbeck & Ronquist 2001, Ronquist *et al.* 2011). Rather than selecting a substitution model a priori, the Bayesian MCMC analysis itself was used to sample across the General Time Reversible (GTR) model space (Huelsenbeck *et al.* 2004). Priors regarding the nucleotide frequencies and substitution rates of the GTR matrix, the shape parameter of the gamma distribution of the rate variation, the proportion of invariable sites and the overall rate were set to vary across the three partitions (the fourth partition, containing the indels information, was combined with the *nad1* partition). To produce an ultrametric, rooted phylogenetic tree, a relaxed Thorne-Kishino 2002 (TK02) clock model, with an underlying strict clock model, was used.

MrBayes ran two simultaneous, independent analyses (two runs), which is the default setting. Each run was run for 500 million generations, with a relative burn-in of 25%. The Monte Carlo chain was sampled every 1000 generations. Per run, four separate chains were used, three of which were heated with a temperature of 0.5. The starting values for the topology and branch lengths priors were derived from an ultrametric tree that was provided at the start of the analysis. *Pelargonium antidysentericum* was set as outgroup.

For a full overview of all the prior settings, see figure 2.1.

```
begin mrbayes;
        set Beagledevice=GPU Beagleprecision=Double;
        set autoclose=no nowarn=yes;
        outgroup antidysentericum214;

 [currently specified groups if you wish to use these]
        charset 179_trnL_F =  1 -  1029;
        charset 137_ITS =  1030 -  1643;
        charset 37_nad =  1644 -  1846;
        partition currentPartition = 3: 179_trnL_F, 137_ITS, 37_nad;
        set partition = currentPartition;
        lset applyto=(1, 2, 3);

        lset nst=mixed rates=gamma;
        unlink statefreq=(all) revmat=(all) shape=(all) pinvar=(all);
        prset applyto=(all) ratepr=variable;
        prset brlenspr=clock:uniform;
        prset clockvarpr=TK02;
        mcmcp ngen= 500000000 relburnin=yes burninfrac=0.25 printfreq=1000
        samplefreq=1000 diagnfreq=10000 temp=0.05 nchains=4 savebrlens=yes Starttree=current;
        startvals tau=gen.22444000 V=gen.22444000;
        mcmc;
end;
```

Figure 2.1. Settings used for the MrBayes 3.2.1 analysis of *Pelargonium*.

Analysing the results

After running for approximately 250 million generations, MrBayes was terminated. To analyse the results, Tracer version 1.5 (Rambaut & Drummond 2007) was used. This program can be used to assess the effective sampling sizes (ESS) of the used priors, and to visualize whether the two separate runs have converged sufficiently. The results can be seen in figure 2.2.

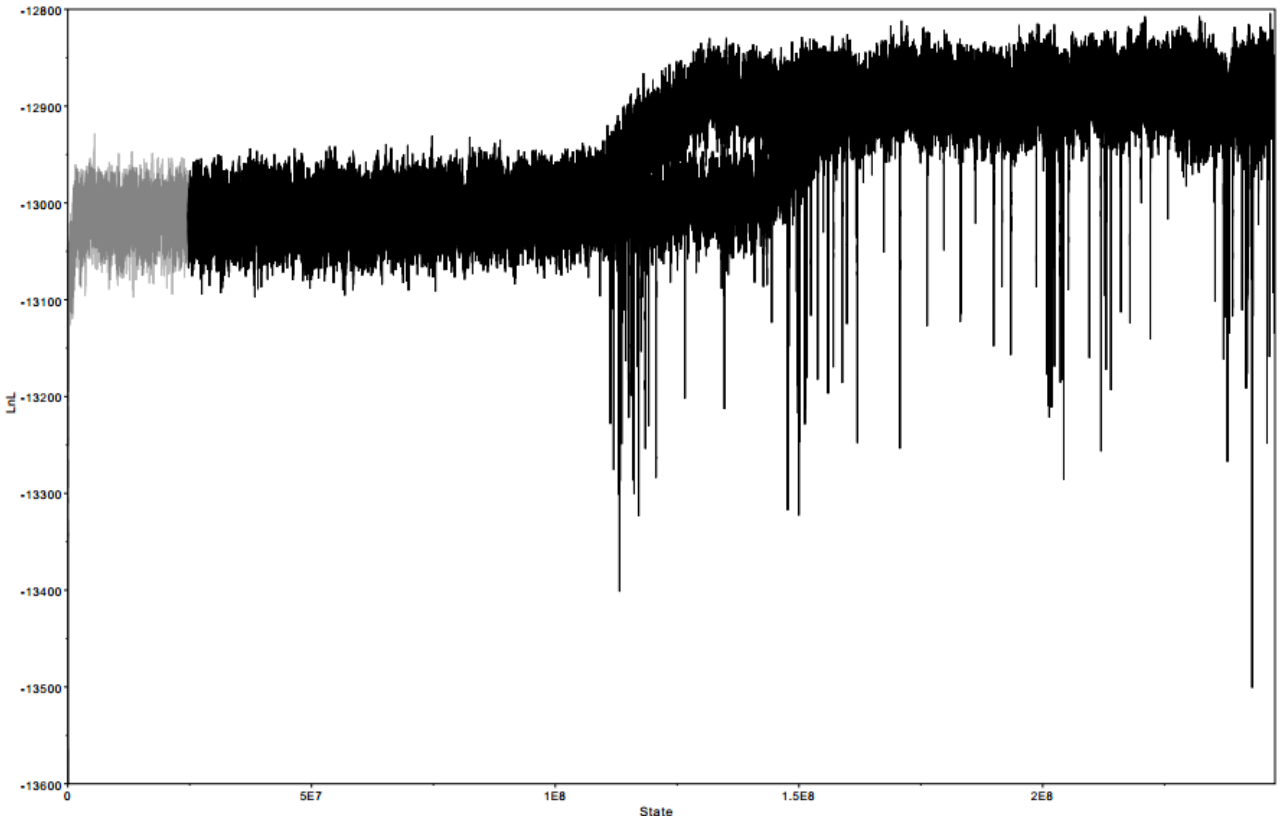

Figure 2.2 Tracer visualization of loglikelihood (LnL) values of the two independent runs of the MrBayes analysis. LnL values range from -13600 to -12800 (y-axis), and the number of generations ranges from 0 to almost 250 million (x-axis).

As can be seen, the two individual runs stabilized and converged after approximately 150 million generations. Therefore, only trees generated in the second part of the analysis (between 150 million and 250 million generations) were used to create a 50% consensus tree, summarizing the two independent runs. In order to do so, the first part of the analysis was removed manually. This was done by opening the two treefiles, containing all the trees generated by each of the two MrBayes runs, in gedit, the text-editor of Linux, and removing the trees from the first generations. As the size of these treefiles was over 9 Gigabyte, they could not be opened by gedit, which only works on files with a size of up to 4 Gigabyte. Therefore, the two treefiles were cut into several smaller files, by using the following function:

```
split -b 500 m treefile.txt treepart
```

where "500 m" defines the size of the output files (in this case, 500 Megabyte), "treefile.txt" is the name of the treefile, and "treepart" is the name of output file (resulting in the files "treepart1.txt", "treepart2.txt", et cetera).

Tracer indicated that the ESS of all parameters of just the last 100 million generations were meeting the requirements. Finally, one consensus tree and 200 trees for future phylogenetic analyses were obtained, all based on the final 100 million generations of MrBayes.

Second analysis

As the results of the first analysis were a bit surprising (see next section), MrBayes was run for a second time, with the same settings and dataset. As the two runs of the second analysis reached divergence after approximately 80 million generations, the analysis was terminated and analysed using Tracer (see figure 2.3). Again, ESS of the parameters were met or exceeded the minimum required size. A consensus tree was created, and 200 random trees were obtained for further analyses.

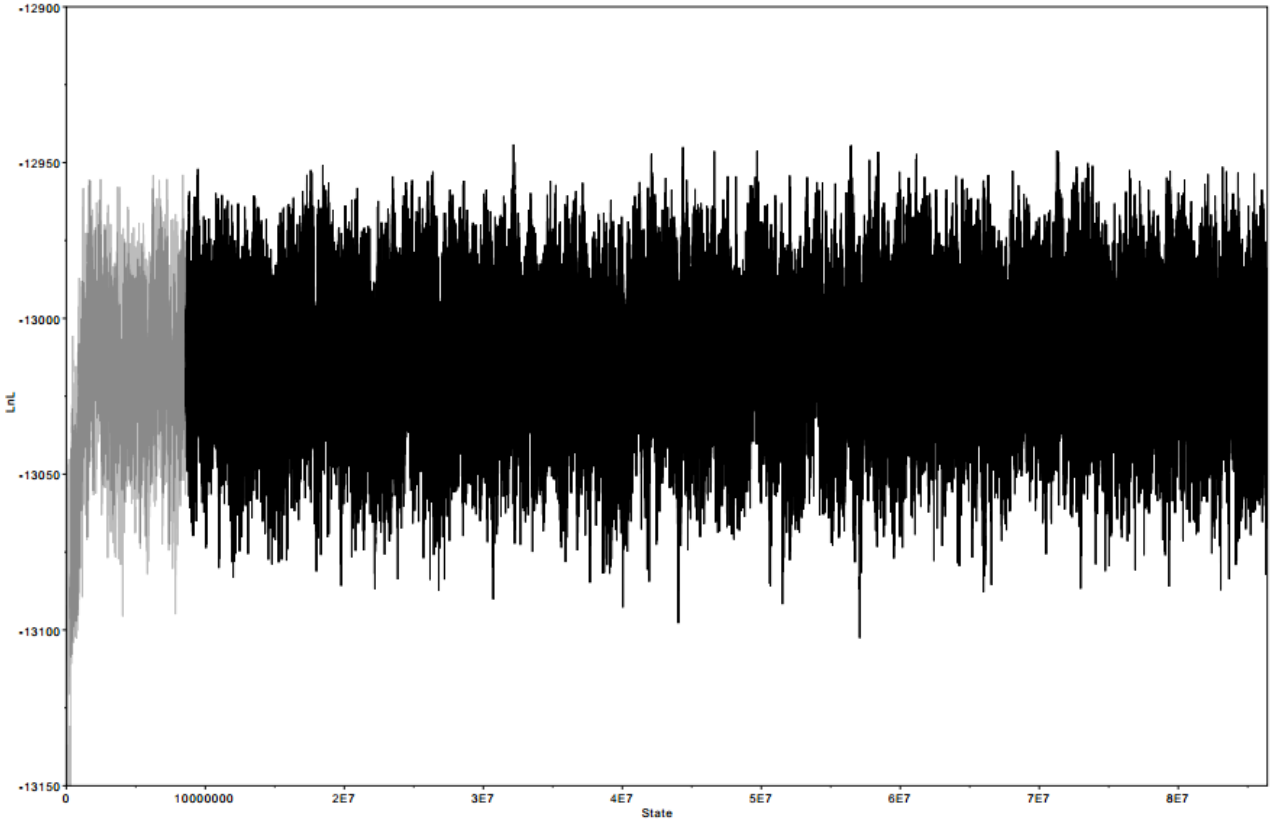

Figure 2.3. Tracer visualization of the second MrBayes analysis. LnL values range from -13150 to -12900 (y-axis), and the number of generations ranges from 0 to over 85 million (x-axis).

Results

The results from the two analyses are shown in figures 2.4 and 2.5, respectively.

As can be seen, the consensus tree derived from the first MrBayes analysis is fundamentally different from the consensus tree of the second analysis. The most striking difference is that in the first tree, clades B and C are sister clades, with clade A as the outgroup, whereas in the second tree clades A and B are sister clades, with clade C as outgroup. In all previous phylogenetic inferences of *Pelargonium*, the topology of the second tree was found, both in older (e.g., Bakker *et al*. 2004, 2005) and very recent studies (Weng *et al.* 2012).
Another remarkable difference is the support for the individual clades, indicated by the posterior probabilities of each clade as calculated by MrBayes. For some nodes in the first tree, these probabilities are significantly lower than in the second tree. The best example of this is the node leading up to the B and C-clades, which has a probability of 0.78. In the second tree, all the major, more basal nodes have a probability of at least 0.99.

As there is no way to distinguish between the two different topologies, all analyses in this thesis have been performed on two samples of trees, one for each topology. From now on, the first topology will be indicated as the A(BC)-topology (as clades B and C are sister clades), and the second topology will be called the (AB)C-topology (as clades A and B are sister clades).

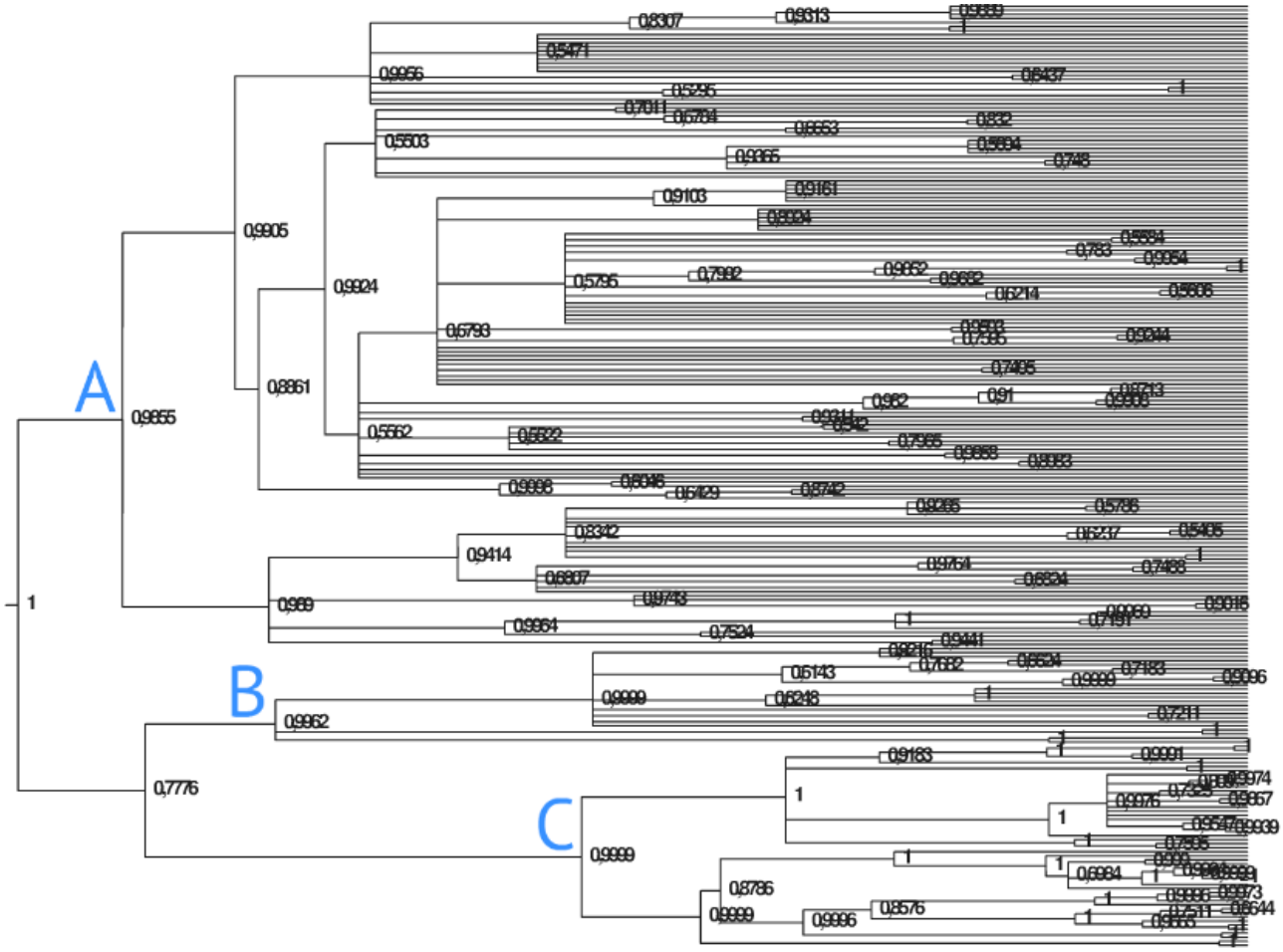

Figure 2.4. 50% Consensus tree of the first MrBayes analysis. Numbers depicted at the nodes correspond with the posterior probability of each clade, as calculated by MrBayes. The three main clades (A, B, and C) are denoted by a letter. To improve readability, the names of the taxa have been removed.

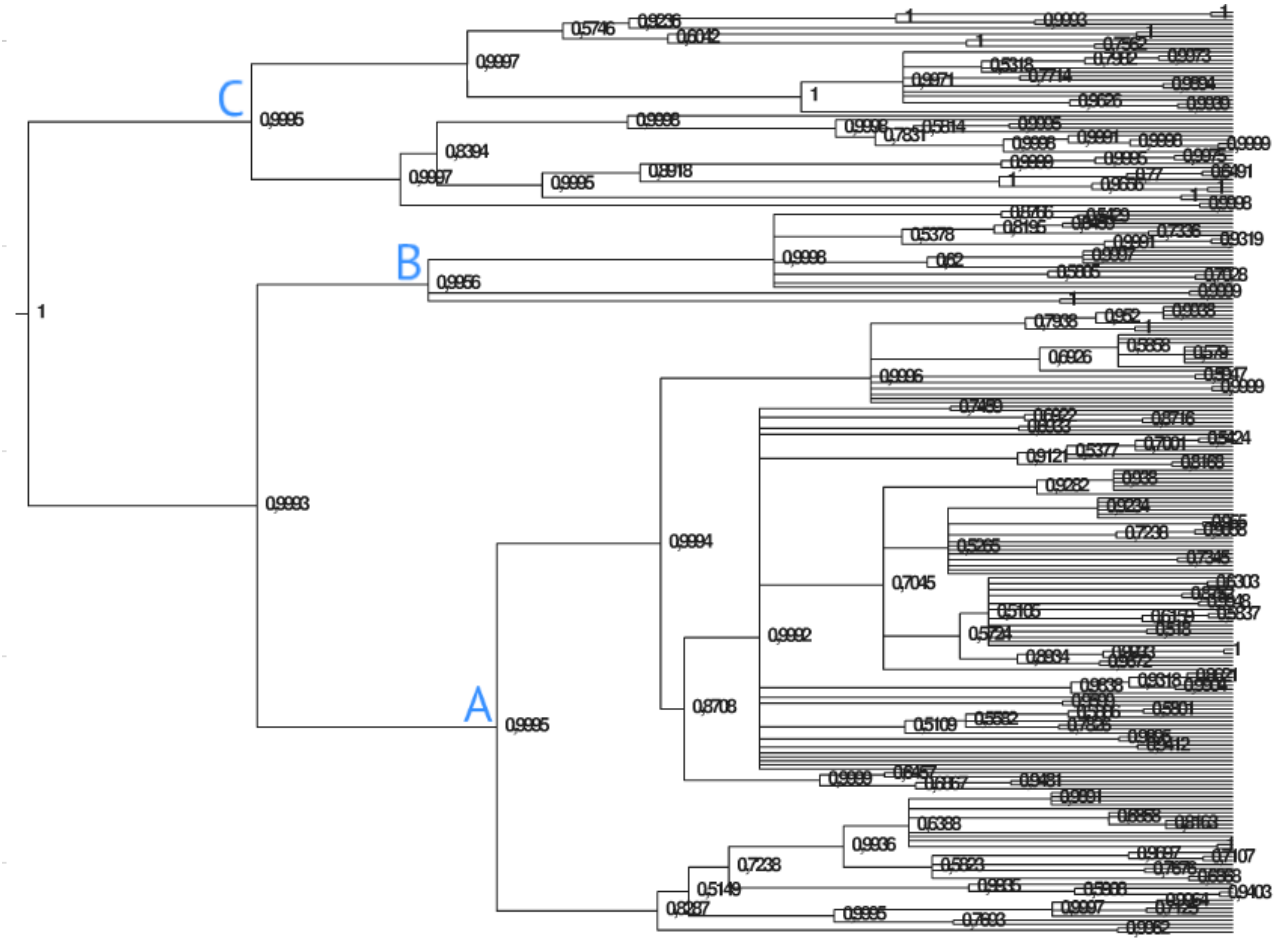

Figure 2.5. 50% Consensus tree of the second MrBayes analysis. Numbers depicted at the nodes correspond with the posterior probability of each clade, as calculated by MrBayes. The three main clades (A, B, and C) are denoted by a letter. To improve readability, the names of the taxa have been removed.

Materials and Methods

Mode of nectar spur evolution

In this thesis, I treat nectar spur length as a continuous character, to be optimised over a set of phylogenetic trees. There are several ways to optimise a continuous character over a phylogenetic tree, but the most used and most straightforward method is Brownian-motion (Felstenstein 1985 and 1988, Pagel 2002), also called the random walk-model. In a standard random walk-model, a character evolves each small instant of time (or genetic change) with a mean change of zero and an unknown constant variance. Furthermore, the rate of change is independent of previous changes in the same branch or in other branches. This means that the random walk-model has four important assumptions: 1) it assumes that the character evolution in a certain part of the tree is not influenced by the evolution in other parts of the tree, 2) that there is no directional tendency in trait evolution over the tree, 3) that character evolution occurs in a gradual rather than punctuational way over time, 4) and also that the rate of change never accelerates or slows down.

The validity of these assumptions can be disputed, as has been done in a number of studies.
The first assumption states that character evolution in a certain branch of the tree is not influenced by the evolution in other branches. This implies that closely related species are not necessarily more similar to each other than to distantly related species. However, the notion that closely related species resemble each other is one of the most fundamental aspects of evolutionary biology. This effect of shared ancestry has been found in sleep time in mammals (Capellini *et al.* 2008), susceptibility to pathogens in plants (Gilbert & Webb 2007), skull evolution in guenons (Cardini & Elton 2008), and a whole range of other studies.
Regarding the second assumption, that there is no directional tendency in character evolution, Whittal and Hodges (2007) found a significant trend towards longer nectar spurs in North American *Aquilegia* flowers. A trend towards larger cranial capacity in Hominids has been described by Pagel in 2002. And in a meta analysis of 89 studies concerning the rate and direction of evolution, Siepielski *et al*. (2009) found that in many of these studies a directional tendency was described.
The third assumption, that character evolution occurs gradually over time, can also be challenged. Nectar spur evolution in *Aquilegia* occurs in quick punctuational bursts rather than in a gradual way (Whittal & Hodges 2007), as does the evolution of the number of caudal-fin rays in Loricariinae catfishes (Covain *et al.* 2007). In a more general study, Venditti & Pagel (2008) describe the occurrence of punctuational character evolution in a large number of cases, including fish morphology, molecular evolution and language evolution.
The last assumption, regarding the tempo of character evolution, can also be disputed. For example, it has been found that in plant lineages with both tropical and temperate species, tropical species have much higher rates of character evolution (Wright *et al.* 2006). Furthermore, a highly important concept in evolutionary biology is the so-called adaptive radiation: the process in which a group of species evolves and radiates rapidly in response to a particular new ecological circumstance, after which evolution slows down. Adaptive radiations have been described in bacteria (Rainey & Travisano 1998), birds and mammals (Hedges *et al.* 1996), and in many other organisms.

Pagel's Transformations

It is quite obvious that, due to these assumptions, the standard random walk-model suffers from some severe shortcomings. These shortcomings can be overcome using several methods and adaptations of the standard random walk-model. I will account for these shortcomings in my thesis by using the method developed by Pagel (1999 and 2002). It consists of several simple adaptations

of standard Brownian-motion, known as Pagel's Transformations. There are four transformations, one for each assumption of the random walk-model. This method is also known as phylogenetic generalised least squares, or PGLS.

In order to use PGLS, the phylogenetic tree first has to be transformed into a variance-covariance matrix, showing the shared evolutionary time. In this matrix, the distance from the root of the tree to the tip for each species (the variance) is shown in the diagonal elements, and the distance from the root to the most common ancestor of each pair of species (the covariance) in the off-diagonal elements (Pagel 1999). See figure 2.6 and table 2.2 for a visual description of this method (figure and table created by Isabella Capellini, see http://nunn.rc.fas.harvard.edu/groups/anthrotree2011/wiki/3507d/Isabella_Capellini__BayesTraits.html). By changing certain elements in the variance-covariance matrix, and testing whether this modified matrix provides a better fit to the data than the original matrix (using likelihood values), the shortcomings of the standard random walk-model can be overcome. Short descriptions follow below, for a full explanation see Pagel (1999 and 2002).

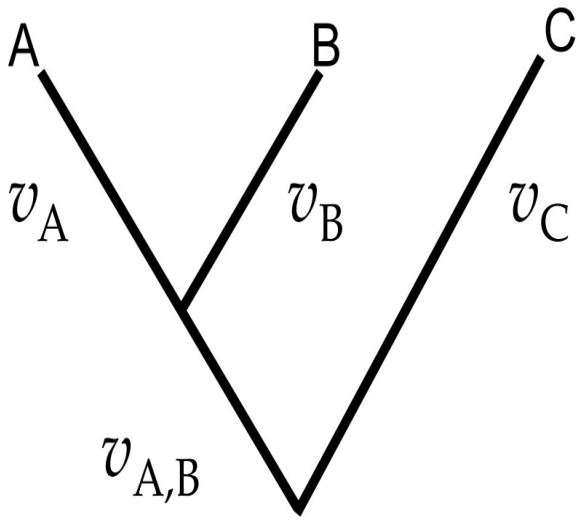

Figure 2.6. Phylogenetic tree with branch-lengths. Created by Isabella Capellini.

	A	B	C
A	$v_A+v_{(A,B)}$	$v_{(A,B)}$	0
B	$v_{(A,B)}$	$v_B+ v_{(A,B)}$	0
C	0	0	v_C

Table 2.2. Variance-covariance matrix corresponding to the phylogenetic tree depicted in figure 2.6. Diagonal elements in the matrix are equal to the distance from the root to the tip for each species, and off-diagonal elements are equal to the distance from the root to the most recent shared common ancestor for each pair of species. Created by Isabella Capellini.

Phylogenetic signal: Pagel's λ

The first of Pagel's Transformations deals with phylogenetic signal: the measure of phylogenetic correlation. It can be defined as the similarity in trait value between species as a result of their shared ancestry. With a high phylogenetic signal, closely related species will resemble each other more than distantly related species. The symbol for phylogenetic signal is (Pagel's) λ.

Pagel's λ can vary between 0 (there is no phylogenetic correlation at all) and 1 (the similarity in trait values between species is directly proportional to the time of shared evolution, one of the basic assumptions of Brownian-motion). In order to test for the presence of a phylogenetic signal, the variance-covariance matrix is transformed; each off-diagonal element in the matrix (the time of shared evolution for each pair of species) is multiplied by λ. As a consequence, only the inner branches of the phylogenetic tree are changed (i.e., shortened), whereas the outer branches, the branches leading up to species, remain unchanged. In the case of a strong phylogenetic signal, hardly any changes are made to the branches (if λ=1, λ*branch length = branch length), causing the

phylogeny to remain unchanged; it already adequately depicts the level of phylogenetic correlation. In the case of a relatively weak phylogenetic signal, the more basal branches are reduced. If there is no phylogenetic correlation at all, the basal branches will collapse (if $\lambda=0$, λ*branch length = 0), resulting in a massive polytomy, a so-called 'star phylogeny' (see figure 2.7).

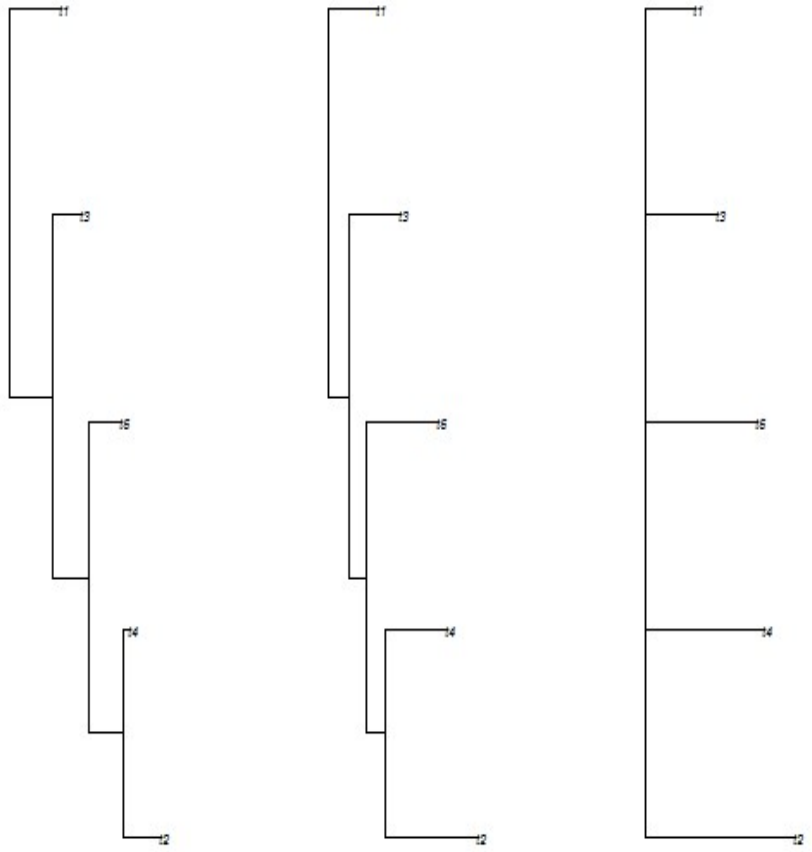

Figure 2.7. Three phylogenetic trees with varying λ-values. The first tree has a λ-value of 1 (i.e., it is unchanged). The second tree has a λ of 0.5, resulting in shorter basal branches. The third three has a λ-value of 0, which causes all the inner branches to collapse, creating a star phylogeny.

Directional random walk: β

The second of Pagel's Transformations deals with directionality in character evolution. In the standard random walk-model, for every instant of time there is an expected change with a mean of zero and a constant variance. With directional random walk, the mean expected change is β, which can be either positive or negative. It is important to realise that β does not transform the variance-

covariance matrix; it works directly on the Brownian motion-formula itself. The value for β is calculated by a regression analysis of species' trait values and total distance from the root of the tree (i.e., the diagonal elements the matrix.) In the case of the existence of a directional trend, species on longer branches, which are species that have diverged more from the root, will tend to have evolved more into a specific direction. Hereby it is important to note that in an ultrametric tree with no fossil data or extinct species, the distance from the root to the tip will be the same for each taxon, since each species is living in the present. Therefore, directionality can not be tested in ultrametric trees. Since I use ultrametric trees for my analyses, I will use a different method to test for the presence of a directional trend (I will use the QuaSSE method, see further in this chapter).

Gradual or punctuational evolution: Pagel's κ

Pagel's third Transformation covers the mode of character evolution: do characters evolve gradually, as is the assumption of the random walk-model, or do they evolve in a punctuational way? The measure of punctuationality is called Pagel's κ. Pagel's κ raises the individual branches of a tree to a power. Similar to β, it does not transform the variance-covariance matrix. κ may vary from 0 to 3, where κ=0 means that evolution is independent of branch length (if κ=0, branch length^κ = 1). This implies that there is a high rate of change in each speciation event, indicating punctuational evolution. On the other hand, if κ=3, long branches stretch far more than short branches, indicating that long branches contribute more to evolution than short branches. This can be seen as evidence for gradual evolution. κ=1 is the default value of κ (if κ=1, branch length^κ = branch length), also indicating gradual evolution (longer branch still contribute more to character evolution than shorter branches). See figure 2.8 for the effect of different values of κ on the branch lengths of a tree.

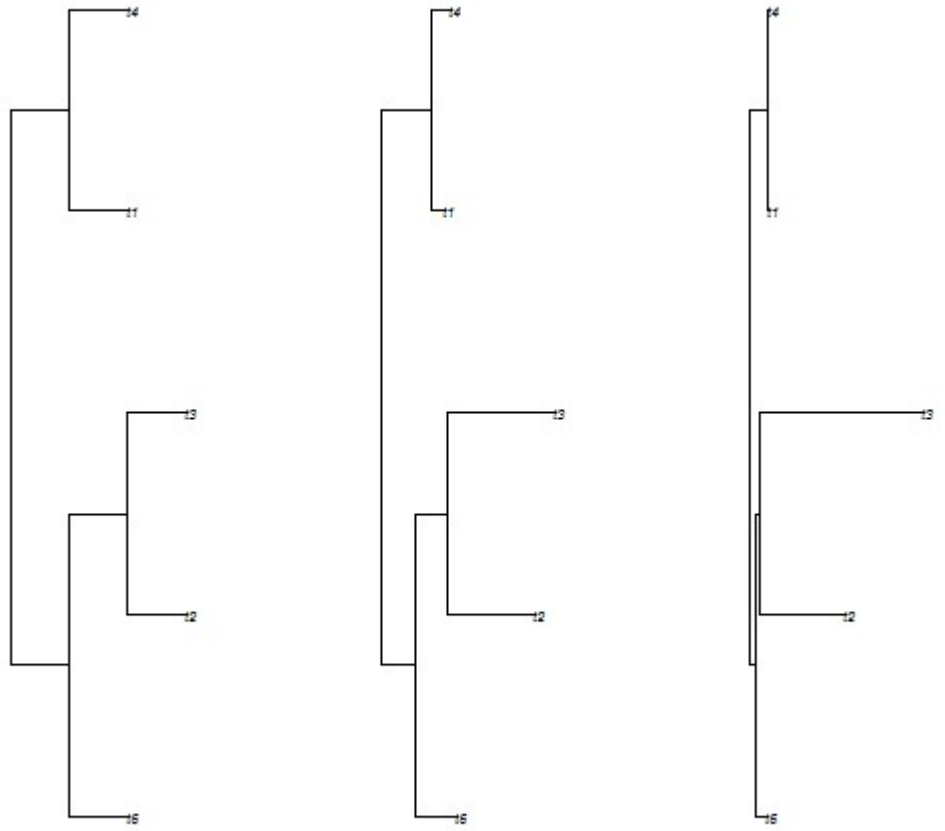

Figure 2.8. Three phylogenetic trees with varying κ-values. The first tree has a κ-value of 0, which has the effect that under Brownian-motion, every branch of the tree contributes equally to character evolution, which indicates punctuational evolution. The second tree has a κ-value of 1 and is the default tree, whereas the third tree has a κ-value of 3. In these two trees, longer branches contribute more to evolution than shorter branches, indicating gradual evolution.

Tempo of evolution: Pagel's δ

The last of Pagel's Transformation deals with the tempo of evolution: does the rate of change remain constant over time (the assumption of Brownian motion), or has it increased or decreased? The measure of this is Pagel's δ, which raises both the diagonal and the off-diagonal elements in the matrix to a power of δ. It can vary between near 0 and 3, where 1 is the default, implying gradual evolution (as assumed in the random walk-model). Since the longest paths in the variance-covariance matrix are the paths leading to species (i.e., the diagonal elements), if δ approaches 0 the longer branches (the outer branches, leading to species) will decrease relatively more than the shorter branches (the more basal branches). This means that shorter paths have contributed more to evolution than longer paths, which can be seen as evidence that the rate of evolution has decreased over time, and may imply adaptive radiation. In that case, character traits changed rapidly at the beginning of evolution, after which character evolution slowed down. However, if δ=3, longer

branches will increase more in length than shorter branches, indicating that the rate of evolution has increased over time. This may imply later species adaptation. See figure 2.9.

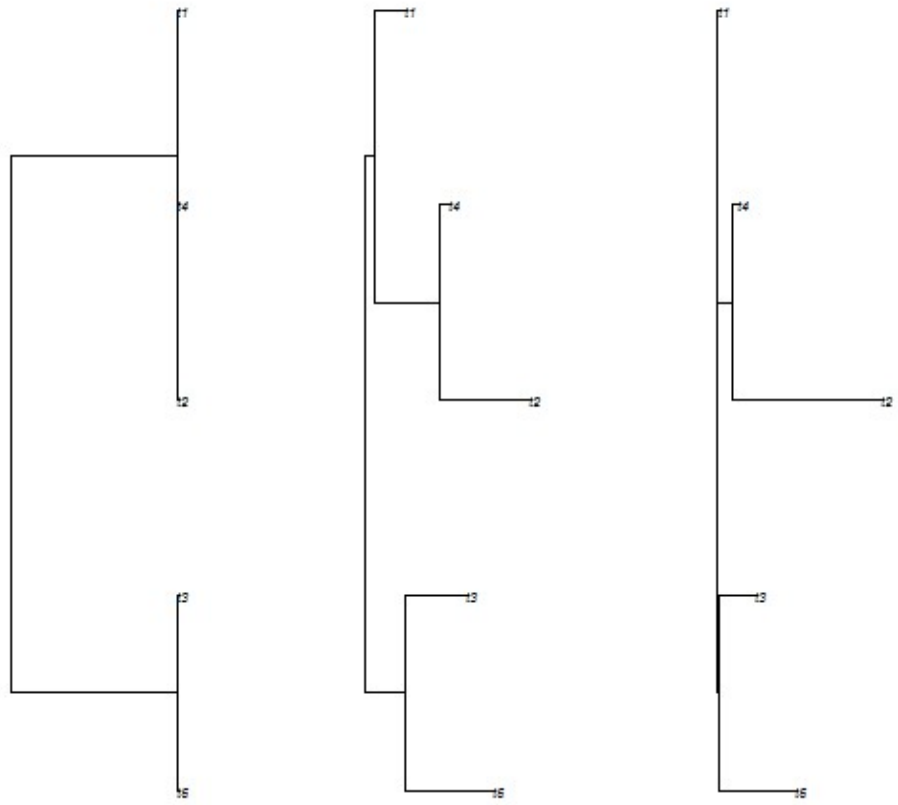

Figure 2.9. Three phylogenetic trees with varying δ-values. The first tree has a δ-value of (almost) 0, which implies a decrease in evolutionary change over time. The second tree has a δ of 1, indicating gradual evolution. The δ-value of the third tree is 3, which means that the rate of evolutionary change actually increased over time.

Pagel's Transformations in R

All of Pagel's Transformations can be calculated using the function "fitContinuous" in the R-library GEIGER (Harmon *et al.* 2009). "fitContinuous" allows one to fit several models of continuous character evolution to a phylogenetic tree and a corresponding character. Available models are, amongst other models, the standard random walk-model as well as all of Pagel's Transformations. By running both a random walk-model as well as, for instance, Pagel's λ, it can be tested whether a random walk-model that incorporates phylogenetic signal provides a better fit to the data than a simple model of Brownian-motion. Comparison of models can be done by a likelihood-test as well as by using Aikaike's Information Criterion (AIC, Aikaike 1974). Apart from a likelihood-value, running the λ-model will also give you a λ-value.

To test if this λ-value is significantly different from λ=0 and λ=1, it is first necessary to transform your phylogenetic tree. In the case of λ, the tree can be transformed using the function "lambdaTree", also available in the GEIGER-package. Other tree-transformations are also possible (e.g., kappaTree and deltaTree). In the case of λ, you need to perform three transformations: one with λ=1, one with λ=0, and one with the λ calculated by "fitContinuous". The next step is to run a standard random walk-model over all of these trees. Comparing the likelihood-values of these three trees calculated by the random walk-model will tell you if the λ-value calculated by fitContinuous is significantly different from λ=0 and λ=1.

In this thesis, I calculated the values for λ, δ and κ for 50 to 200 trees of both the (AB)C-topology and the A(BC)-topology, using the method I described above. As Hardy noted in 2006, and as was seen in chapter 1, the use of median or mean character values may not necessarily be more accurate than the use of, say, minimum or maximum character values. In my analyses I therefore used median, minimum and maximum spur length, and also the range and the variance of spur length for each species. Furthermore, λ, δ and κ-values were calculated for each of the major clades (A, B, and C).

Ancestral nectar spur length estimation

Given a phylogeny that accurately describes the shared ancestry of a collection of species, and information about certain character states or values in these species, it is possible to make predictions about the character values in the ancestors of the current species (Pagel 2002). If done correctly, this could provide us with valuable insights in the processes that shaped the evolution of these species (Ronquist 2004, Pagel *et al*. 2004). Question that may be addressed by ancestral character estimation include whether ancestral characters affect speciation rates, which characters are more strongly correlated with cladogenic events, and in which sequence ancestral characters evolved (Hardy 2006). Ancestral characters have been reconstructed in a wide range of studies. Examples include the inference of ancient fruit characters in Cornaceae (Xiang & Thomas 2008), the reconstruction of hormone receptors (Thornton *et al.* 2003), the design of HIV-vaccines (Gaschen *et al.* 2002), the reconstruction of ancestral ecologies (Hardy 2006) and even the analysis of ancestral behaviours (Schwarz *et al.* 2003).

Methods and assumptions

There are several ways to reconstruct ancestral characters (Ronquist 2004, Hardy 2006, Xiang & Thomas 2008). However, almost all available methods have two basic assumptions: 1) the assumption that there is a rate constancy, and 2) the assumption that branch lengths give us information about the amount of character evolution (Ekman *et al.* 2008). The first assumption states that the rate of character change is a constant in any part of the tree. However, Schluter (1997), Cunningham (1998), and their respective colleagues conclude that the rate of character change can vary widely across the tree (see also further in this chapter). The validity of this assumption may therefore be called questionable at most. The second assumption implies that branch lengths in a phylogenetic tree, whether they are a measure of time or genetic change, convey information about the amount of morphological or ecological evolution (depending on the character used). This assumption can be disputed as well. It is definitely possible that branch lengths contain information about character evolution, but it also known that in many cases, there is no clear correlation between genetic change (branch length) and morphological change (character evolution) (Ekman *et al.* 2008). A good example is adaptive radiation, where much morphological evolution occurs in relatively little time (Cunningham 1999).

To account for the possibility that branch lengths may in fact not contribute any information about character change, one of the ancestral character estimation methods I used in this thesis was squared-change parsimony. Simple, quick and elegant, an important feature of a parsimony analysis is that branch lengths have no effect on the ancestral character estimations (Ekman *et al.* 2008). A parsimony approach finds the ancestral character reconstruction that minimizes the evolutionary costs; the explanation that requires the smallest number of changes or the smallest amount of evolution (i.e., the most parsimonious explanation) is deemed the most likely explanation (Ronquist 2004). Squared-change parsimony means that the cost for changing from state x to state y will be squared: $(x-y)^2$. The consequence of squaring the difference is that the costs for transitions from short spurs to longer spurs are the same as the costs for changes from long spurs to short spurs (e.g., $(6-4)^2 = (4-6)^2 = 4$). Squared-change parsimony also implies that 'big' changes (going from the shortest spurs to the longest, or vice versa) are less likely than smaller changes (e.g., $(1-3)^2 = 4 < (1-8)^2 = 49$). A squared-change parsimony analysis can be performed quite easily in R using the package APE (Paradis *et al.* 2004). This can be done by first setting all the branch lengths to 1 using the function "compute.brlen", and afterwards ancestral characters can be estimated using the function "ace". More detailed information will be provided further on in this chapter.

Maximum likelihood

Quick and simple as it may be, a parsimony analysis does have certain disadvantages. The most important shortcoming is that, even though the found ancestral character reconstruction may be the most parsimonious, that does not necessarily mean it accurately portrays evolutionary history (Ronquist 2004). In fact, several other reconstructions may be almost as likely as the parsimony construction, but will be ignored by this method (Frumhoff & Reeve 1994). A way to work around this problem is to use (maximum) likelihood as reconstruction method, which does use branch lengths to infer ancestral characters. A likelihood analysis will incorporate some model of character evolution to infer ancestral characters, which may not necessarily provide the same results as a parsimony analysis. For continuous characters, the model used in likelihood analyses is Brownian-motion or random walk (Felsenstein 1985 and 1988). Just as the parsimony reconstruction, a likelihood analysis can be performed in the R-library APE.

However, a likelihood analysis using a random walk-model has two serious shortcomings: 1) the basic assumptions of a Brownian motion-model of evolution may be disputed, and 2) a likelihood analysis treats the provided phylogenetic tree as accurately portraying evolutionary history; likelihood does not take phylogenetic uncertainty into account. Both these problems can be dealt with by using the R-libraries APE and GEIGER (Harmon *et al.* 2009).

Model choice in R

The first shortcoming is that the basic assumptions of random walk-models is questionable. This was already discussed in the previous section, so only a quick summary will be given here. A random walk-model of evolution assumes that character evolution is independent of evolution in other parts of the tree, has no directional bias, occurs gradually and never speeds up or slows down. As discussed before, in many situations one or more of these assumptions may prove to be invalid. To test this, the function "fitContinuous", of R-package GEIGER, will be used. "fitContinuous" fits several models of continuous character evolution to a phylogenetic tree and corresponding character. Apart from the models already addressed before (i.e., standard Brownian-motion and Pagel's Transformation), there are four additional possibilities: an Ornstein-Uhlenbeck model, an early burst model, a white noise model, and a trend model.

The Ornstein-Uhlenbeck model (Butler & King 2004) is an adaptation of the random walk-model. It fits a model of Brownian-motion with a central tendency towards an optimal character value. Using this model, you can account for the possibility that evolution is not random, but that it is in fact driven to an optimal character value. The early burst model (Harmon *et al*., paper in revision), also called the ACDC model, fits a model where the rate of character change can increase or decrease over time. The white noise model assumes that there is no phylogenetic signal, and that all species values are drawn from the same normal distribution. The trend model fits a random walk-model with a trend, either positive or negative. As trends in character evolution can only be detecting used non-ultrametric trees (see previous section), I have not taken this model in consideration.

Using "fitContinuous", a total of seven models was fitted to the phylogenetic tree and the nectar spur lengths. The most likely model was chosen based on a likelihood ratio test and Aikaike's Information Criterion. The next step was to transform the tree to fit the predictions of the best model. This was done by using the "deltaTree"-function of GEIGER. The branch lengths were transformed to fit the model of evolution, and then ancestral characters were estimated using the function "ace". Two methods were used for the reconstruction: maximum likelihood (ML) and restricted maximum likelihood (REML). According to the online description of the "ace"-function (see http://127.0.0.1:29993/library/ape/html/ace.html), it was shown that when using ML the variance of the Brownian-motion process is estimated with a downward bias, while there is no such bias when using REML. To compare the results of ML and REML, I will use both methods. The output of ace consists of calculated ancestral character values, their 95% credibility intervals, and the estimated Brownian-motion rate parameter.

Phylogenetic uncertainty

The second problem of using using maximum likelihood (and also parsimony) to infer ancestral character values on a phylogenetic tree, is the possibility that this tree does not accurately portray evolutionary history. In fact, it is even highly likely that one single tree will never be the "true tree" (Huelsenbeck 2000). The solution to this problem is fairly straightforward: in my analyses I used a set of trees, rather than one single tree. I performed two separate analyses, one for each topology (A(BC) and (AB)C), and I used 200 randomly-chosen trees for each analysis. The important nodes (e.g., the nodes leading up to the major clades) in the trees were identified using the function "oldest.mrca" from the R-package Phytools (Revell 2012). For each of these nodes, the mean reconstructed character value (and its corresponding 95% credibility interval) was calculated based on the 200 individual reconstructions.

In the scientific literature, considerable debate exists on the reliance of the different character estimation methods (e.g., Ronquist 2004, Hardy 2006, Ekman *et al.* 2008, Xiang & Thomas 2008). In this analysis, the three methods used (squared-change parsimony, maximum likelihood and restricted maximum likelihood) were compared for the most important nodes. The differences between the optimization methods were also visualized by calculating the sum of the squared-differences between two different optimizations for each node in tree. This was repeated for every tree that was used.

All three optimization methods were used on 200 trees of each topology, with minimum, median, and maximum nectar spur lengths.

A possible problem with the methods described above is that by averaging the results over a range of trees, the reconstructed nectar spur value for the most ancestral node in the tree will always be

the average of all nectar spur values in the tree (P. Linder, pers. comm.). To test whether this affects the results, I divided the nectar spur lengths randomly over all species in the tree and reconstructed the value for the most ancestral node for 200 trees of the (AB)C-topology. I repeated this analysis four times.

Stochastic character mapping

As the ancestral character estimation methods in R may suffer from some shortcomings (see previous section), a different method was used as well: stochastic character mapping (Huelsenbeck *et al.* 2003), as implemented in SIMMAP (version 1.5, Bollback 2006). Rather than reconstructing character values at individual nodes, stochastic character mapping, also known as (Bayesian) mutational mapping, maps character changes directly along the branches of a phylogeny. A SIMMAP analysis results in a complete overview of all the changes between character states, as well as the average dwelling time per character state. Additionally, SIMMAP allows ancestral state reconstruction for each individual node in the phylogeny by calculating the marginal posterior probability of each possible character state (Bollback 2006). The program has been used in dozens of studies, for example to analyse the egg-deposition behaviour in darters (Kelley *et al.* 2012), the evolution of venom proteins in reptiles (Casewell *et al.* 2012), and leaf-shape variation in *Pelargonium* (Jones *et al.* 2009).

SIMMAP only works with discrete characters, with a maximum of seven states per character. Therefore, median nectar spur length was divided into seven discrete states (see table 2.3). State ordering was set to linear, indicating that nectar spur length in state 1 is larger than nectar spur length in state 0, et cetera.

State number	Spur length (in cm)
0	$x = 0$
1	$0 < x \leq 1$
2	$1 < x \leq 2.3$
3	$2.3 < x \leq 3.6$
4	$3.6 < x \leq 4.9$
5	$4.9 < x \leq 6.2$
6	$6.2 < x$

Table 2.3. Character states of median nectar spur length as used in SIMMAP.

An important part of a SIMMAP analysis is the specification of the two priors: the bias parameter prior and the rate parameter prior. For characters with more than two states, only two bias priors are available: an empirical prior and an equal (1/k) prior, the latter of which was selected. For the rate parameter a gamma distribution prior was chosen. The gamma distribution is characterised by the α, β, and κ-parameters, which were set at 44.156, 0.439, and 60, respectively. These settings were based on an MCMC run of the dataset (using 100000 cycles, a sampling frequency of 200, and a burnin of 1000), which was analysed using an R-script provided on the SIMMAP website (Bollback 2009).

Several sampling settings need to be specified before the analysis can be started. The number of samples was set to 50 and the number of prior draws to 1 (as recommended). As 201 trees were used (200 randomly chosen trees plus 1 50% consensus tree), this resulted in 10050 individual mutational maps (201 * 50 = 10050). Ideally one would require a larger number of samples, but due to unknown technical problems, SIMMAP automatically shut down when setting the amount of

samples higher than the current value. To check whether 10050 mutational maps is sufficient, SIMMAP was run four times, and the results were compared. If the output of the four independent runs are similar, this would indicate the four runs all have converged on the same results, implying that a higher number of samples would not have had any effects on the final outcome.

SIMMAP was run using median nectar spur lengths and 201 trees of the (AB)C-topology.

Analysis of lineage diversification rate shifts

To test for the presence or absence of evolutionary shifts in lineage diversification rates within the *Pelargonium* phylogeny, I used the MEDUSA method, developed by Alfaro, Brown and colleagues (Alfaro *et al.* 2009, Brown *et al.* 2012). MEDUSA is a function contained in the library GEIGER (Harmon *et al.* 2009), part of the open statistical computer environment R. MEDUSA is an acronym for Modelling Evolutionary Diversification Using Stepwise AIC, which is an accurate description of this function. It fits a series of increasingly complex birth-death models to an ultrametic phylogenetic tree. Whether these models are retained or discarded is based on Aikaike's Information Criterion (AIC, Aikaike 1974). MEDUSA has been used to analyse diversification rate shifts in dozens of studies and systems, such as in coprinellus mushrooms (Nagy *et al.* 2012), salamanders (Rabosky & Adams 2012), and spiny-rayed fish (Bannikov & Carnevale 2011).

MEDUSA works in a simple yet elegant way. First it fits a straightforward birth-death model to a phylogenetic tree. In this first model, it is assumed that both speciation (birth) and extinction (death) rates are the constant in the whole phylogeny. Then, the likelihood of this particular combination of phylogenetic relationships and clade ages and sizes, given the maximum-likelihood values for birth- and death-rates, is calculated. As a final step, the AIC score is calculated with a simple formula: $AIC = 2k - 2\ln L$, where lnL is the likelihood value and k is the number of parameters needed to describe the model. This simple model, with which the MEDUSA analysis starts, has just two parameters; one for the birth-rate and one for the death-rate.
Now a more complex model is fitted to the data. In this model, the presence of a single breakpoint somewhere in the phylogeny is tested. Branches originating in this breakpoint have different birth-death rates than the branches leading up to this point. Again, likelihood-value and AIC score are calculated, with one important difference: this more complex model now has five parameters: two birth-rates (b1 and b2), two death-rates (d1 and d2), and one breakpoint. If this new model has an AIC score which is four units lower than the previous model (which is the threshold for a significant increase in model fit (Burnham & Anderson 2003)), it is considered to be a better description of the phylogeny than the previous model, and it is retained. Now an even more complex model is fitted to the data, and so on. This is process is continued until no improvement in AIC score can be found, or until the maximum number of breakpoints is found (the default limit is twenty breakpoints). After this forward selection process where increasingly complex models are adopted, a backwards elimination procedure is now performed, to test whether a simpler model may have a better AIC score. Once both these procedures have been finished, the model that best fits the data is retained, and information about breakpoints and birth- and death-rates can be recovered. For a full description of the MEDUSA method, see Alfaro *et al.* (2009).

During the MEDUSA analysis, there are several choices to be made regarding model-type, parameter-choice and calculation-limits. However, one of the most important decisions that has to be made is what kind of phylogenetic tree will be used. There are two options: one can either use the normal, unaltered tree, or a so-called skeleton tree. A skeleton tree is a tree from which most taxa have been removed, so only one taxon remains per major clade. To use such a skeleton tree, a table with species richness data per clade has to be added; the result is a tree where each clade

consists of one taxon, and the number of species per clade can be found in a separate table (see figure 2.10 and table 2.4). The alternative is to offer the full phylogeny to MEDUSA. The biggest difference between these approaches is the fact that when using the skeleton tree, breakpoints can

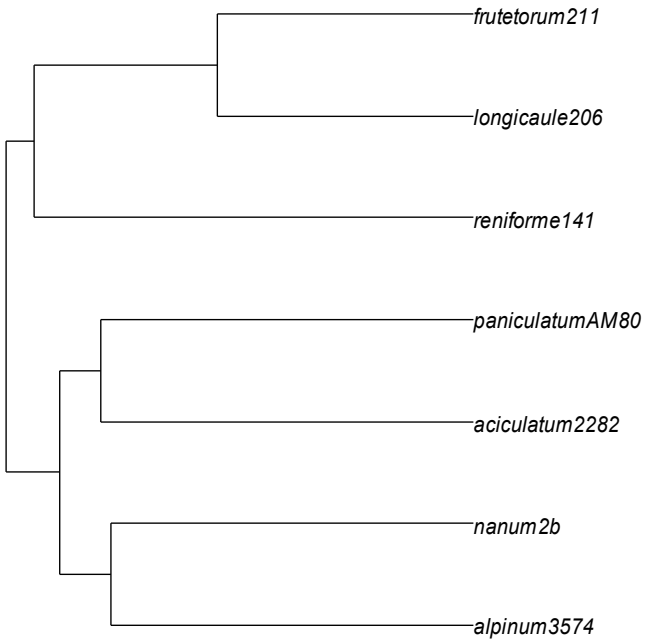

Figure 2.10. Skeleton tree of the (AB)C-phylogeny, to be used in a MEDUSA analysis. Per major clade only one taxon remains.

Taxon	Sections	Number of species
aciculatum2282	Hoarea, Ligularia, Magnistiupulaceae	87
paniculatumAM80	Polyactium, Otidia, Cortusina	25
alpinum3574	Pelargonium, Campylia	44
nanum2b	Nanum	1
reniforme141	Peristera, Reniformia	31
frutetorum211	Ciconium, Subsucculentia	25
longicaule206	Myrrhidium, Chorisma, Jenkinsonia	24

Table 2.4. Species richness information per major clade in the *Pelargonium* phylogeny. The number of species per clade is shown. The species linked to each clade correspond with the species depicted in figure 2.10.

only be found in the branch leading up to a particular clade. When the full tree is used, on the other hand, breakpoints can also be found within a clade. To test for the differences between these two approaches, I performed a MEDUSA analysis both on the full tree as well as on the skeleton tree.

The MEDUSA analysis was run with 200 trees of the (AB)C-phylogeny and 200 trees of the A(BC)-phylogeny. The default settings were used. This includes, amongst other settings, a maximum number of twenty breakpoints, the possible use of both pure-birth as well as birth-death models, and the comparison of different models using AIC corrected for small sample sizes, AICc.

To make sure the sampling frequency was the same for each clade in the trees, a few species (14 in total) were removed from the trees before the analyses were run, in order to minimise the effects of sampling bias. Furthermore, all double species, subspecies, and varieties were removed from the tree.

The output of the analysis consists of the number and locations of the breakpoints that were found, as well as the corresponding model parameters. Of particular interest are the r-value (= birth-rate minus death-rate) and the value for epsilon (= death-rate divided by birth-rate). The results are plotted onto a tree, so the presence and location of breakpoints can be easily assessed.

Spur length-dependent speciation rates

One of the principle questions in evolutionary biology is whether certain traits or aspects of species can actively influence the speciation or extinction rates of these species; this process is called species selection (Rabosky & McCune 2009). For a long time it has been remarkably difficult, if not impossible, to test whether certain species have higher speciation rates than others, but this has changed with the recent spectacular rise of available molecular data and phylogenetic methods. Testing for the presence of species selection is now possible for binary characters (Maddison *et al.* 2007, FitzJohn *et al.* 2009), qualitative characters, and quantitative characters (FitzJohn 2010). As the spurs lengths which I am interested in are continuous characters, I have used this last method, which is called QuASSE: Quantitative State Speciation and Extinction (FitzJohn 2010). It is implemented in the R-package Diversitree (FitzJohn *et al.* 2012). Diversitree has been used in a number of studies; the analysis of the effect of dispersal ability on passerine birds (Claramunt *et al*. 2011), the effects of the loss of sexual recombination on the speciation of primroses (Johnson *et al.* 2011), and the diversification of animals as diverse as dipsadine snakes (Burbrink *et al.* 2012) and ruminants (Cantalapiedra *et al*. 2011).

QuaSSE

QuaSSE uses likelihood equations to calculate the probability of a specific phylogenetic tree, a distribution of character states, and a model of cladogenesis (FitzJohn 2010). The model used is a birth-death model where the speciation and extinction rates may vary, based on the evolution of a quantitative character. This character is assumed to evolve under a diffusion model. The character can have four different effects on the speciation and/or extinction rates, which are depicted in figure 2.11.

If there is no effect of the character value on speciation or extinction rates, the speciation function will remain constant (blue line in figure 2.11). This serves as the null model, against which other functions can be tested. If speciation/extinction rate and character value are directly correlated, the speciation function will be linear (red line). The two remaining options are a sigmoid speciation function (green line) and a parabolic function (orange line). Of course, the graphical parameters describing these functions may vary, so the speciation and extinction functions may have any possible slope, width, curve and direction. QuaSSE will calculate the likelihood value of the combination of the tree, characters and each speciation and extinction function with its particular parameters. The function that best describes the data can be selected based on its AIC score. A chi-squared test can be used to examine whether a specific function yields significantly different results compared to the null model.

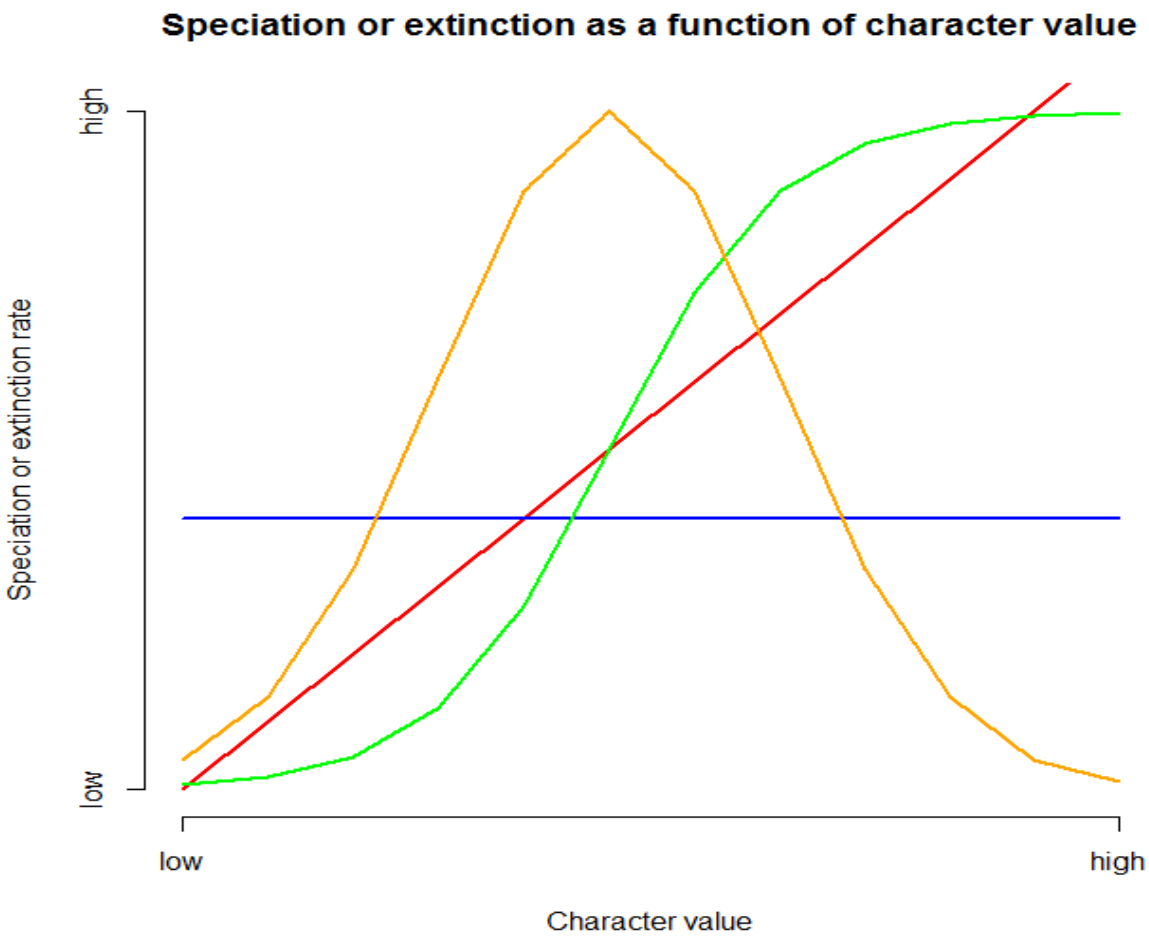

Figure 2.11. Four different shapes of speciation or extinction rate as a function of the character value. The blue line indicates a constant response (no effect of speciation or extinction), the red line is a linear response, the green line a sigmoid, and the orange line a parabolic response.

Rate shifts and directionality

Another function of QuaSSE is that it allows different parts of the phylogenetic tree to evolve with different speciation/extinction rates. A possible implementation for this function is suggested by FitzJohn (2010); use MEDUSA (Alfaro *et al*. 2009, see previous section) to identify the location of a diversification rate shift in the tree, and use QuaSSE to test whether the implementation of this rate shift in the model will result in a better fit of the model. To give an example: FitzJohn (2010) found that in certain clades of primates, body size was positively correlated with speciation, whereas in other clades, the correlation was negative.

A third function of QuaSSE is that allows testing for the presence of an evolutionary trend towards larger or smaller character values. As discussed before, testing for the presence of such a trend is not possible with Pagel's Transformations when using ultrametric trees. With QuaSSE, directionality in ultrametric trees poses no problem. However, the power to detect such a trend is low, and there is "essentially no power to detect the presence of the directional tendency where it

reinforced species selection" (FitzJohn 2010).

Assumptions and shortcomings

Elegant as it may be, QuaSSE is not perfect; it does suffer some limitations (FitzJohn 2010). An important one is that it allows for no interaction between lineages in three. The extinction rate in one part of the tree can not be influenced by the number of species in any other part of the tree. As a consequence, density-dependent evolution or frequency-dependent evolution can not be modelled in QuaSSE.

Furthermore, even though QuaSSE allows speciation and extinction rates to be modelled, there are severe theoretical and practical shortcomings of estimating extinction rates based purely on a tree with no fossil data or extinct species (Rabosky 2010). As there is a complete lack of fossil data in the Geraniaceae (Bakker *et al*. 2005), I have only used QuaSSE to estimate speciation functions.

Settings

I have run QuaSSE on a total of 35 trees of both topologies, using median, minimum, maximum spur length, and the range of spur length. Additionally, as MEDUSA did find some support for the presence of a rate shift in the (AB)C-topology (see results section of this chapter), I used QuaSSE to model two separate speciation functions in this topology in 15 trees. The location of this rate shift was put at the base of the A-clades. Before running QuaSSE, all double species, subspecies and varieties were removed from the tree.

With a few exceptions, QuaSSE was run with all the default options. To account for incomplete sampling in the tree, the sampling frequency was set to 0.72. The lower and upper limit of possible spur length were set to 0 and 10 centimetres, respectively. Standard deviations of spur length, needed for QuaSSE, were calculated based on the variance of spur length previously calculated. Speciation functions with and without drift were fit to the data, to test for the presence of a directional trend.
Apart from these small exceptions, default options were used, as can be found in the Diversitree tutorial on the Diversitree website (http://www.zoology.ubc.ca/prog/diversitree/).

Spur length evolution rate shifts

One of the basic assumptions of the Brownian-motion model of character evolution, is that the rate of character change is assumed to be constant in every region of the tree. We have already seen that this assumption may be violated, and Pagel's Transformations can be used to test whether rates of change are constant throughout the phylogenetic tree or not. However, Pagel's Transformations do not provide any information about the direction and the location of possible shifts in the rate of character change. This is rather unfortunate, considering how the identification and localization of such shifts has been the focus of evolutionary biologists since Darwin (1859, for example see also Simpson 1994 and Estes & Arnold 2007). Over the years, methods for identifying such shifts have become much more sophisticated. However, the most recently developed models require that possible locations for character rate shifts are identified a priori (O'Meara *et al.* 2006), or can only identify one shift (Revel *et al.* 2012). A recently developed method does not have the shortcomings of either of these models. The method auteur (Eastman *et al.* 2011), Accommodating Uncertainty in Trait Evolutiong Using R, contained in the eponymous R-package auteur (Eastman *et al.* 2012), can identify the probability of several shifts of rate change without any a priori information about the location of these shifts. As auteur is a relatively new method, it has not yet been used in any peer-

reviewed studies. However, the authors of auteur did use it to find several highly elevated rates of character evolution in turtles and monkeys (Eastman *et al.* 2011).

Auteur assumes that character evolution occurs along a phylogenetic tree according to a random walk-model. In a method similar to MEDUSA (Alfaro *et al.* 2009), auteur then fits a series of increasingly complex models to the combination of phylogenetic tree and continuous character. Unlike MEDUSA, however, auteur identifies shifts in the rate of change of character evolution, rather than in the rates of diversification and extinction. Another important difference is that auteur uses reversible-jump Markov Chain Monte Carlo (Metropolis *et al*. 1953, Hastings 1970) for its analyses. Different Brownian-motion models, with different positions and directions of rate shifts, are sampled according to their posterior possibility (Bartolucci *et al.* 2006).

The output of auteur, consisting of marginalized distributions of relative rates for each branch in the tree, and corresponding information about the location and direction of probable shifts in rate change, can be assessed using the Java-based software Tracer (Rambaut & Drummond 2007), or using the R-library CODA (Plummer *et al.* 2006). These methods can be used to check whether the effective sample sizes (ESS) of the different rate parameters are large enough. A commonly accepted ESS value is typically larger than 100, to ensure proper chain mixing.

Auteur was run using the function "rjmcmc.bm", with two independent chains, 1000000 generations, a sample frequency of 100, and with a random model complexity as starting point. For the remaining parameters, default options were chosen. Auteur was run on five trees of both topologies (A(BC) and (AB)C). A few species (14) were removed from the trees to ensure equal sampling frequencies in each clade.

Results

Pagel's Transformations

Pagel's λ

Pagel's λ deals with phylogenetic autocorrelation; the measure of similarity in species' trait values as a result of their shared ancestry. Pagel's λ-transformation was fitted to 200 randomly-chosen trees from the A(BC)-topology and the median nectar spur lengths. For every single tree, the λ-model was a highly significant improvement ($p < 0.001$) of the standard random walk-model according to the likelihood test. The average λ-value is 0.69, which indicates a moderately strong phylogenetic signal (λ ranges from 0 to 1)(Pagel 1999). See figure 2.12 for a visualization of the λ-values (plot created using the function "sm.density.compare" from the sm-library (Bowman & Azzalini 2010) in R).

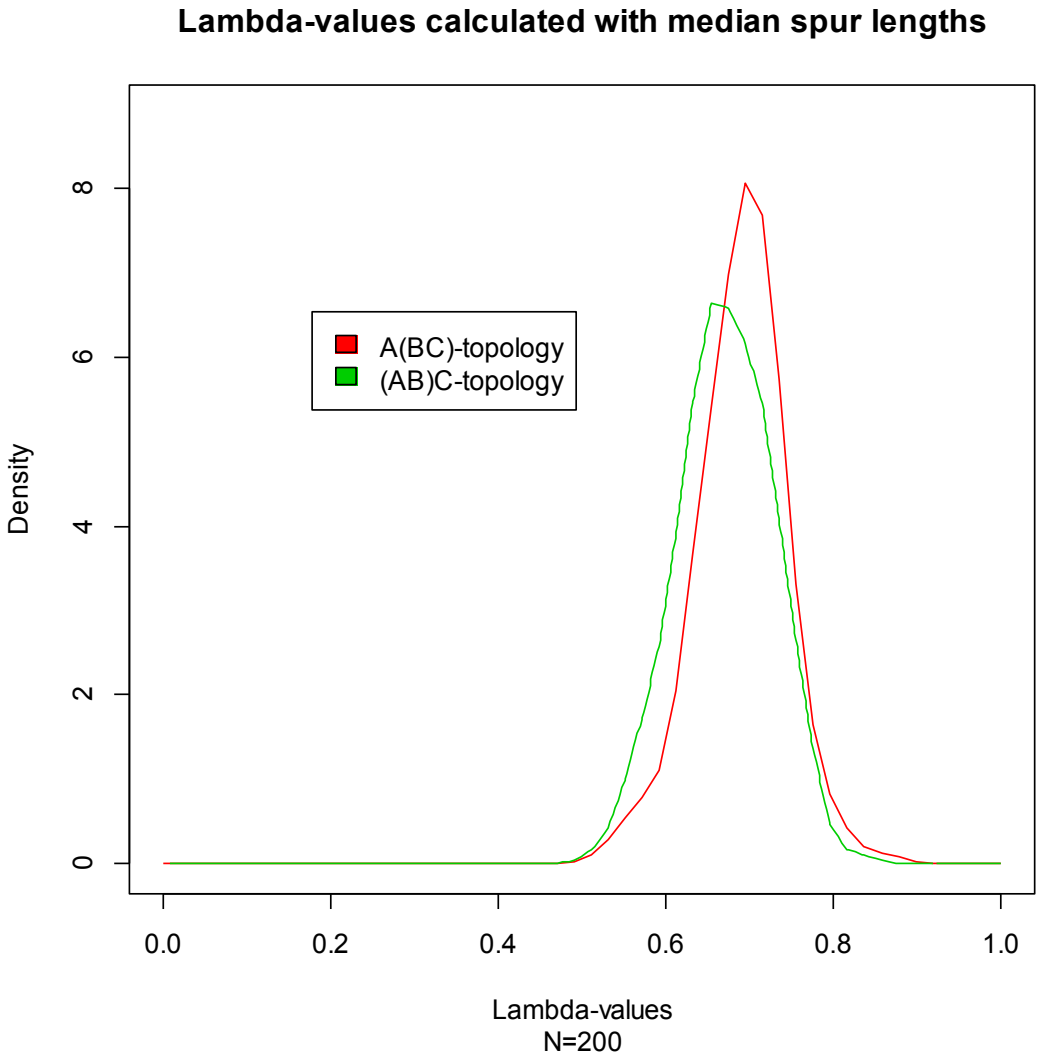

Figure 2.12. λ-values for 200 randomly-chosen trees from both topologies.

According to the likelihood-test, the λ-value calculated by fitContinuous is significantly different ($p < 0.001$) from both λ=1 and λ=0, and according to the AIC-test, the calculated λ-value always provided a better fit than a model using λ=1 or λ=0.

For the (AB)C-topology, the results are highly similar. For every used tree (200), the λ-model was a highly significant improvement ($p < 0.001$). The average λ-value is 0.67 (see figure 2.12) and according to both likelihood-test as well as AIC, this value for λ was a improvement over λ=1 or λ=0.

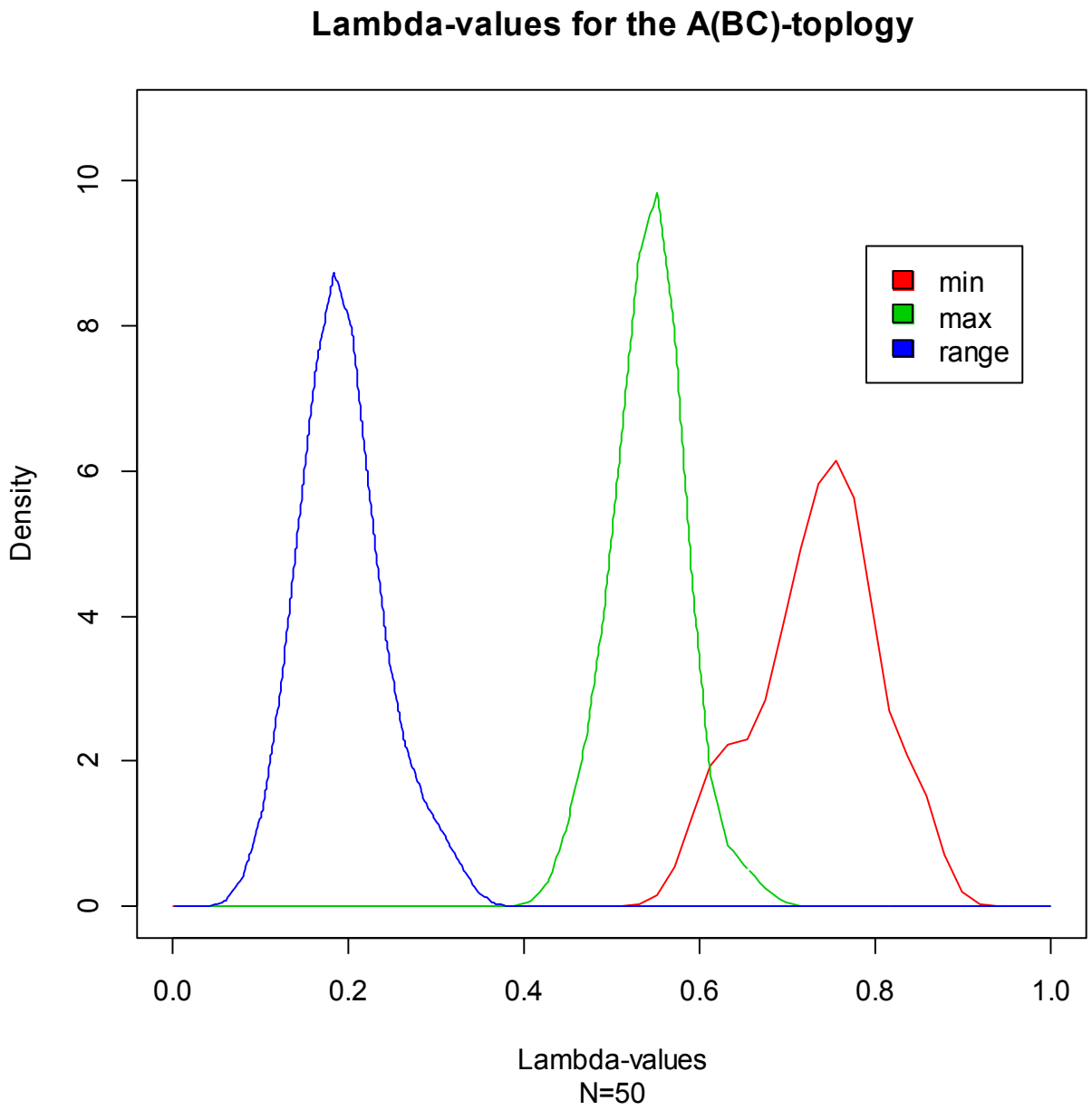

Figure 2.13. λ-values for 50 trees of the A(BC)-topology and different measures of spur length. Minimum and maximum spur length have the highest measure of phylogenetic signal, whereas the range and especially the variance (not shown) hardly have any indiscernible phylogenetic signal at all.

For both topologies, I also fitted a λ-model using minimum and maximum spur lengths, as well as the variance and range of the spur length distribution, to 50 randomly-chosen trees. For the A(BC)-topology, the results can be seen in figure 2.13 (for graphical reasons, the λ-values of the variance have been left out of the plot). As can be seen, minimum spur length has the strongest phylogenetic signal, followed by maximum spur length. The range of spur length hardly has any phylogenetic signal, and λ-value of the variance is indiscernible from λ=0 ($p > 0.05$), which indicates that there is

no phylogenetic signal at all.

For the A(BC)-topology, the average λ-value for minimum spur length is 0.73, which is significantly different from both λ=1 and λ=0 ($p < 0.001$). For maximum spur length, the average λ is 0.54, which is again significantly different from both standard λ-values ($p < 0.001$). The phylogenetic signal of the spur length range, however, with an average λ of 0.19, is only significantly different from λ=0 in 8 out of 50 trees ($p < 0.05$).

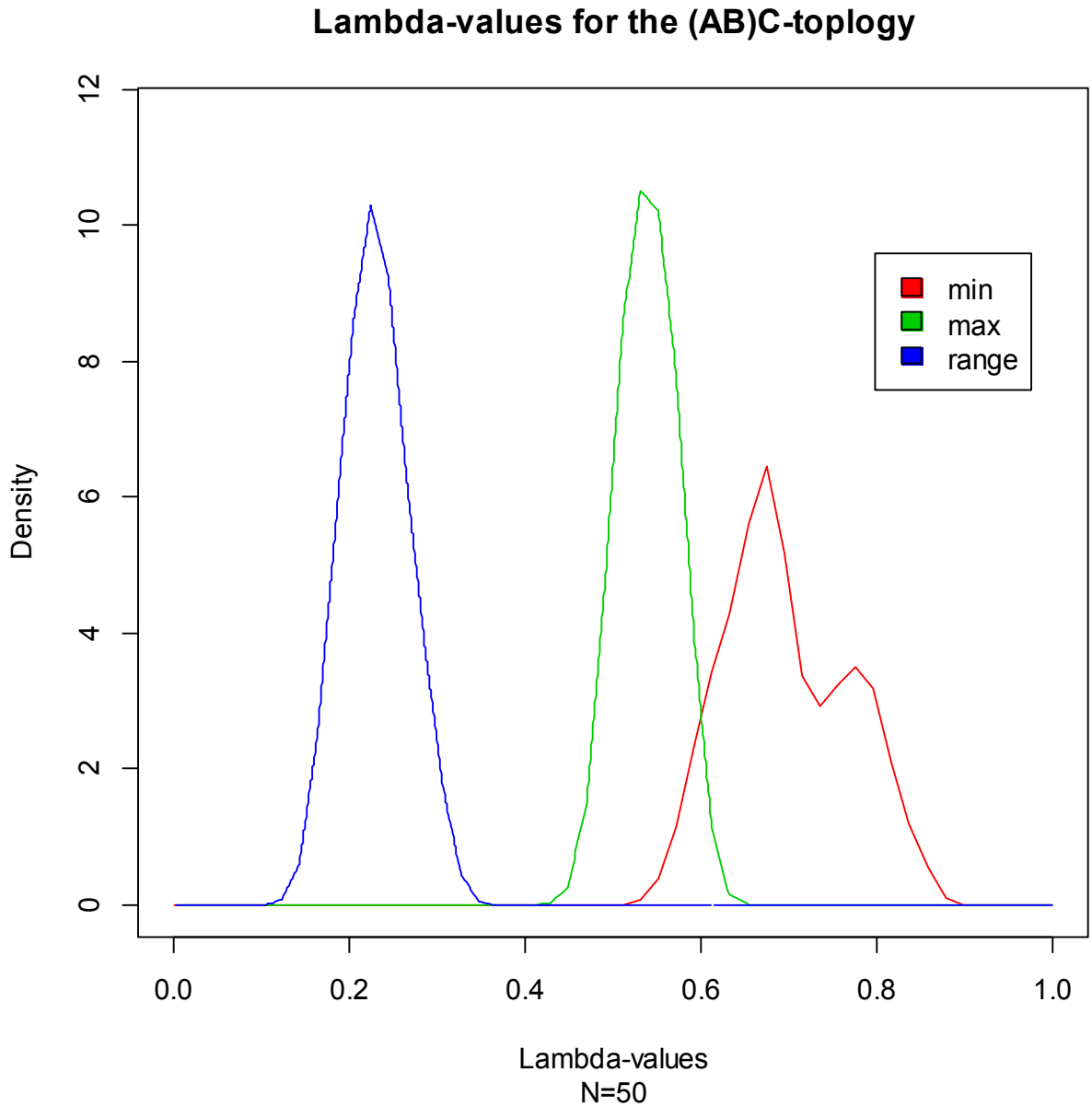

Figure 2.14. λ-values for 50 trees of the (AB)C-topology and different measures of spur length. Minimum and maximum spur length have the highest measure of phylogenetic signal, whereas the range has a relatively low phylogenetic signal.

The results for the (AB)C-topology are similar to the A(BC)-topology (see figure 2.14 (again, λ-values for the variance have been left out of the plot)). The average λ-value for minimum spur length is 0.69 (significantly different from both λ=1 and λ=0 ($p < 0.001$)), indicating a moderately strong phylogenetic signal, and the average λ for maximum spur length is 0.53 (different from λ=1 and λ=0 ($p < 0.001$)). The λ-values for variance and range are, just like in the other topology, much

closer to zero; for range, the average is 0.23, which is only different from λ=0 in 7 out of 50 trees, and for variance, λ is indiscernible from 0, indicating once again that there is no phylogenetic signal in the variance of nectar spur length.

Values for Pagel's Transformations were also calculated by using one of the three major clades, rather than the full phylogenetic tree. Each analysis was performed on 50 trees using the median of spur length (see figure 2.15). The average λ-value for the A-clade of the A(BC)-topology, containing the sections *Hoarea*, *Ligularia*, *Magnistipulaceae*, *Polyactium*, *Otidia*, *Cortusina*, *Pelargonium*, *Campylia*, and *Nanum*, is 0.68. This provides a significant different fit ($p < 0.001$) than a λ-value of either 0 or 1, and according to AIC provides the best model fit. The B-clade, containing sections *Peristera* and *Reniformia*, has an average λ-value of 0.57, which is only significantly different ($p < 0.05$) from $\lambda = 0$, for 13 trees, but is an improvement over $\lambda = 0$ or 1 according to AIC. The C-clade, with sections *Ciconium*, *Subsucculentia*, *Myrrhidium*, *Chorisma*, and *Jenkinsonia*, has an average λ-value of 0.40. This is significantly different ($p < 0.01$) from the standard λ-values, and a model using the calculated λ-value better describes the character evolution than a model using one of the standard λ-values. For each clade the phylogenetic signal is moderately strong.

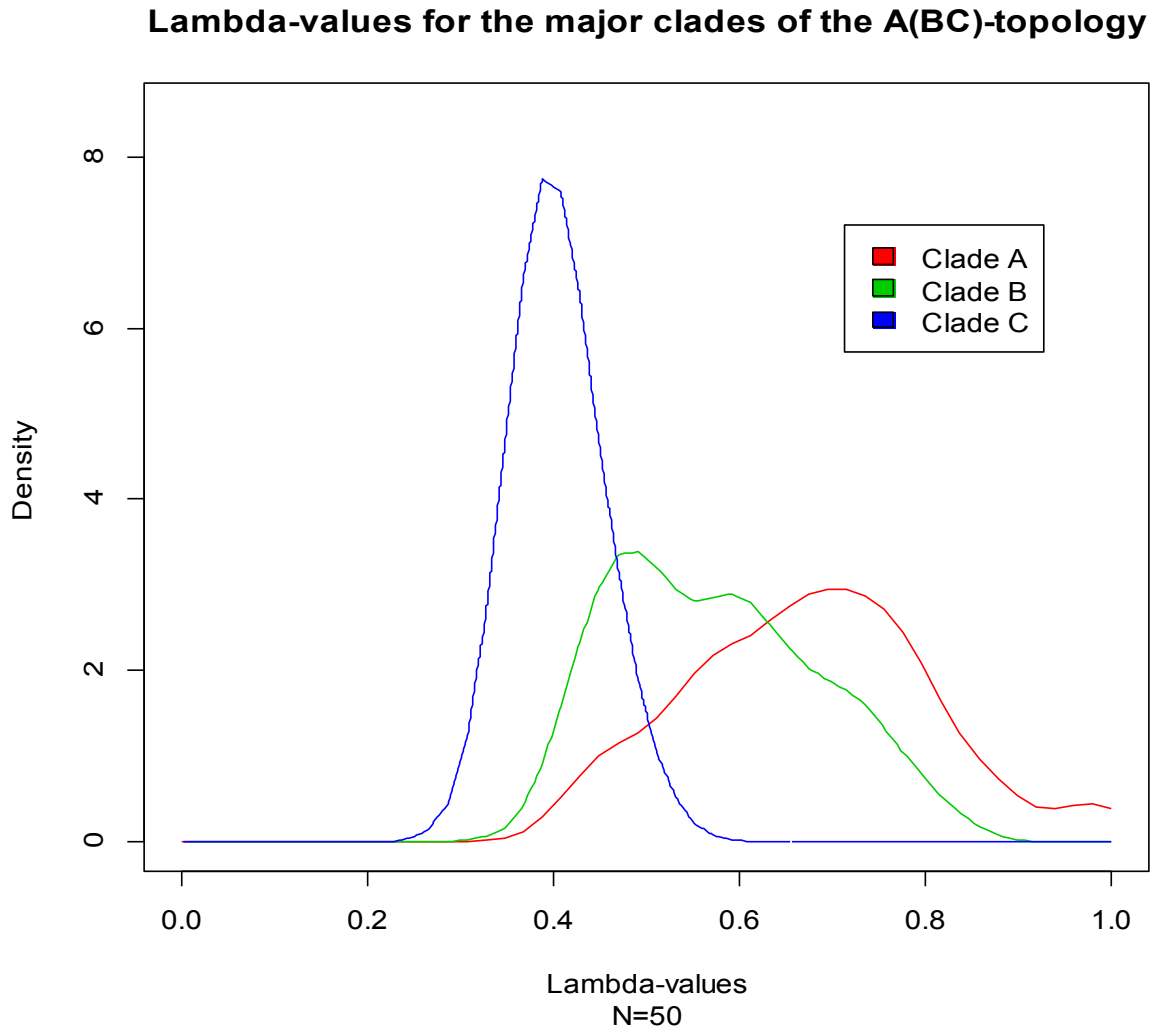

Figure 2.15. λ-values for the major clades of 50 trees of the A(BC)-topology.

The results for the (AB)C-topology are fairly similar to the results for the A(BC)-topology (see figure 2.16). Once again, there is a moderately strong phylogenetic signal for each major clade. The average λ-value for the A-clade is 0.65, for the B-clade 0.60, and for the C-clade it is 0.42. Again, for the A- and C-clade this model provides the best description of character evolution and is significantly different from a model using a standard λ-value (with a p-value of $p < 0.001$ and $p < 0.01$) respectively. For the B-clade however, the calculated λ-value is only significantly different from $\lambda = 0$ for 10 trees. A model using the calculated the λ-value provides the best fit for 47 trees.

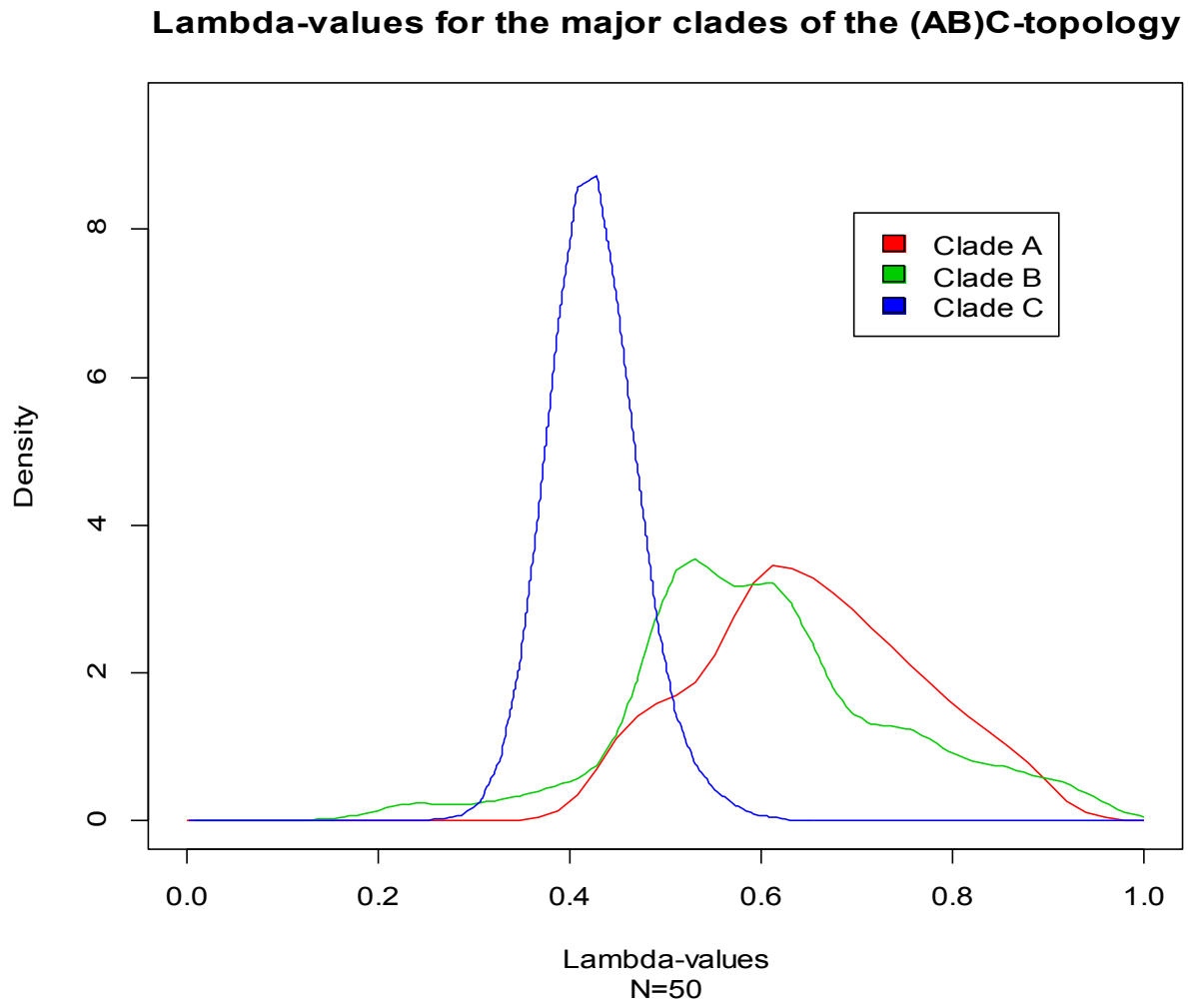

Figure 2.16. λ-values for the major clades of 50 trees of the (AB)C-topology.

Pagel's κ

Pagel's κ ranges from 0 to 3, where a κ-value of 0 means that character evolution is independent of branch length, indicating that evolution occurs in a punctuational way, whereas a κ of 3 means that longer branches contribute more to evolution than shorter branches, which can be seen as evidence for gradual evolution (Pagel 1999). The default value of κ is 1, which indicates relatively mild gradual evolution.
For both topologies, I fitted a κ-model of evolution to 50 randomly-chosen trees. Once again, I used median, minimum, and maximum spur length values, as well as the range and the variance.

The results for the κ-values calculated with median spur lengths can be seen in figure 2.17. The average κ for the A(BC)-topology is 0.35, which is highly significant from ($p < 0.001$), and proves a better fit compared to, $\kappa = 0$, 1 or 3 for every tree. For the (AB)C-topology, the average value of κ is quite a bit lower: 0.22. This value for κ is only different from $\kappa=0$ in 31 out of 50 trees (p-value differs between $p < 0.05$ and $p < 0.001$). However, according to the AIC test, this κ-value does provide a better fit than $\kappa = 0$ (or 1 and 3) in every tree. These are strong indications that longer branches do not necessarily contribute more to character evolution than shorter branches; this means that spur length evolution occurs in a punctuational way, implying that there is a lot of spur length evolution at each speciation event.

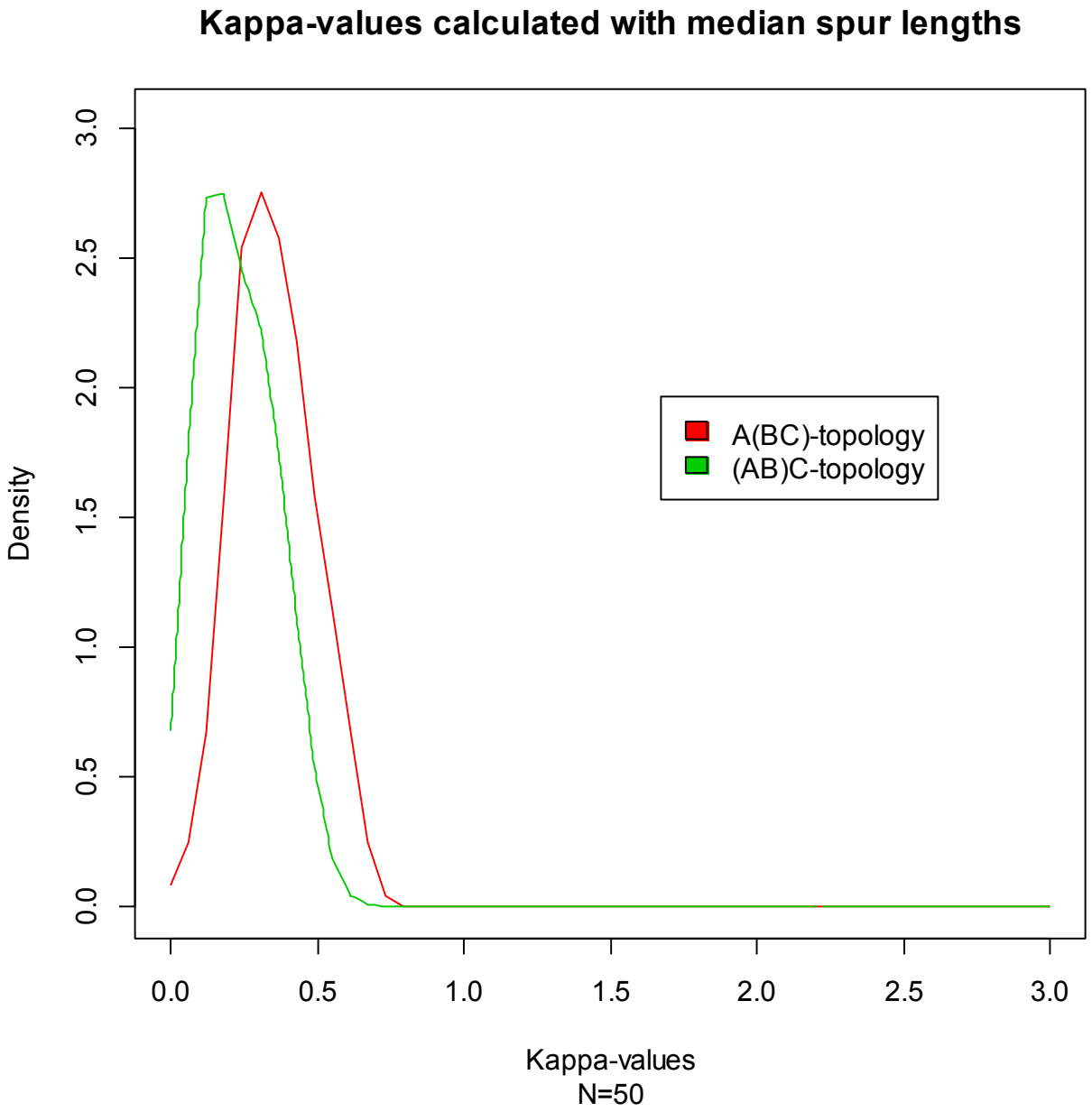

Figure 2.17. κ-values for 50 randomly-chosen trees from both topologies.

As with the λ-values, I also fitted a κ-model of evolution to minimum and maximum spur length, as well as to the range and the variance, for both topologies. The results can be seen in figures 2.18

and 2.19.

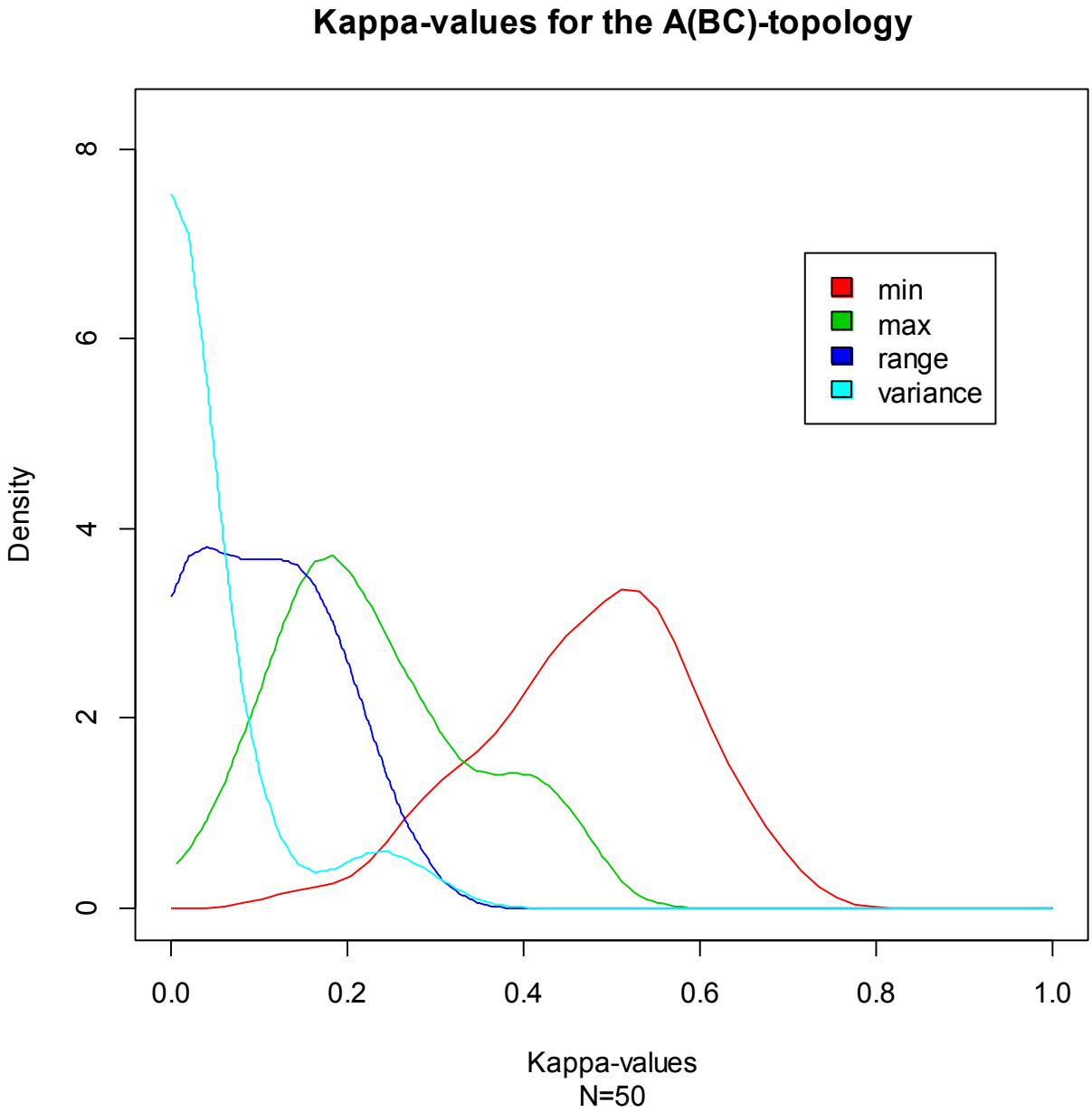

Figure 2.18. κ-values for 50 trees of the A(BC)-topology and different measures of spur length. As can be seen, the highest κ-values of minimum spur length approach the default value of 1, whereas maximum spur length, range and especially the variance of spur length are closer to κ = 0. For graphical reasons, the x-axis (marking the κ-value) ranges from 0 to 1, not to the maximum κ-value of 3.

For the A(BC)-topology, the average κ-value for minimum spur length is 0.47, which is significantly different from κ = 0, 1 or 3 for every tree ($p < 0.001$). According to AIC, a model with the calculated κ-value is also an improvement over a model using one of the standard κ-values. The average κ-value for maximum spur length is remarkable lower: 0.22. This indicated that maximum spur length evolves in a more punctuational way that minimum spur length. The calculated κ-value is always a better fit than a standard κ-value, but is only significantly different from κ = 0 in 32 out

of 50 trees (p-value ranging from $p < 0.05$ to $p < 0.001$).
As with the λ-transformations, the κ-values of the variance and range of spur length are even lower. For the range, the average κ-value is 0.10, which is only significantly different from $\kappa = 0$ in 6 trees ($p < 0.05$). According to AIC, a model using a calculated κ-value provides an improvement over a model using $\kappa = 0$ in 41 out of 50 trees. The average κ-value for the variance of spur length is 0.03. This value is indiscernible from $\kappa = 0$ according to both the likelihood test and AIC.

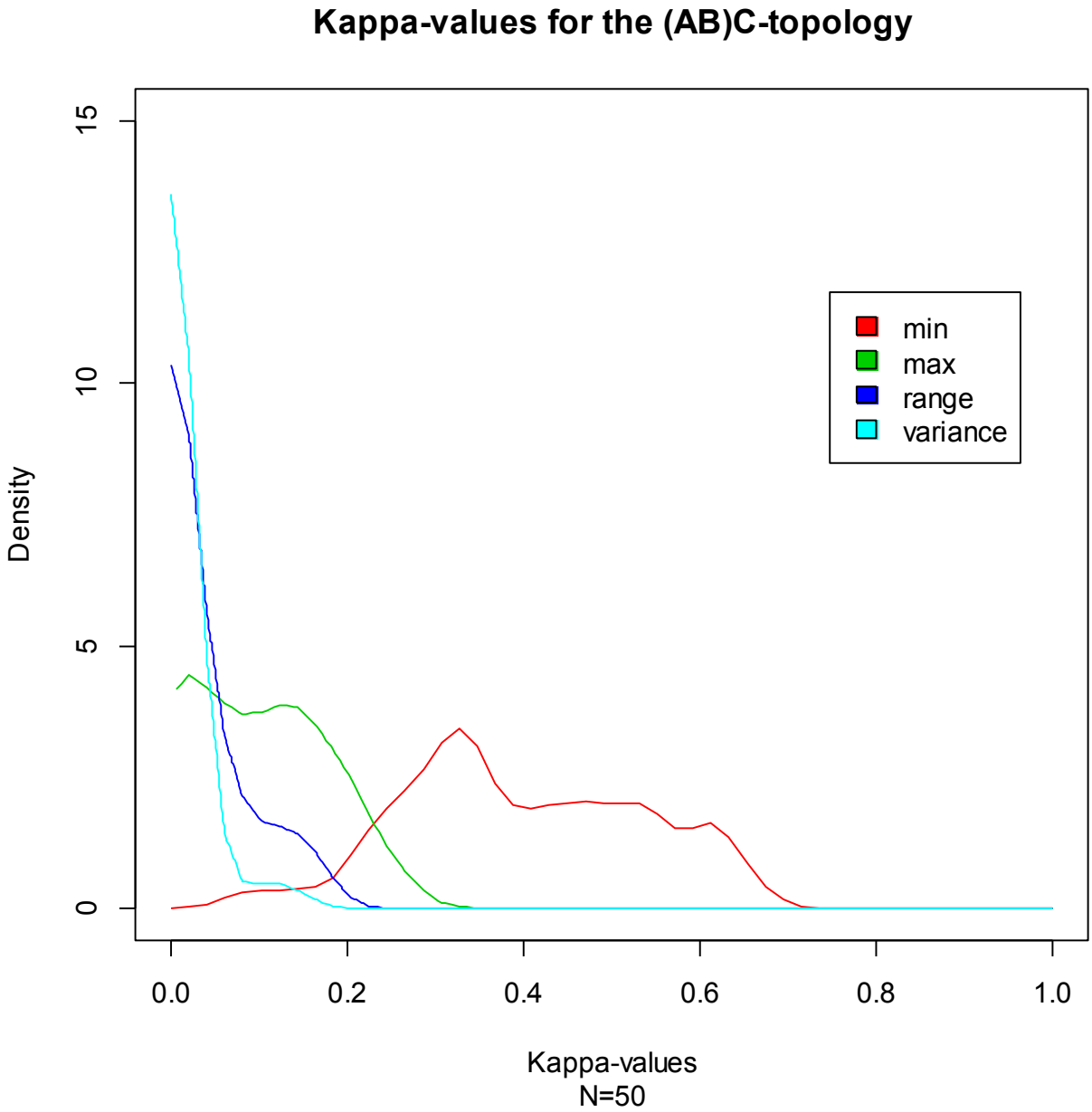

Figure 2.19. κ-values for 50 trees of the (AB)C-topology and different measures of spur length. Similar to the results from the A(BC)-topology, κ-values for minimum spur length are the highest, closely followed by values for maximum spur length. The values for the range and variance approach 0. For graphical reasons, the x-axis (marking the κ-value) ranges from 0 to 1, not to the maximum κ-value of 3.

The results for the (AB)C-topology, shown in figure 2.19, are relatively similar, albeit a bit lower, to

the results of the A(BC)-topology. The average κ for minimum spur length is 0.40, significantly different ($p < 0.001$) to, and a better model-fit than, κ-values of 0, 1, or 3 for every tree. Maximum spur length has an average κ-value of 0.10, which is only significantly different from $\kappa = 0$ in 5 trees ($p < 0.05$), but provides a better model-fit in 40 trees. The average κ-values for the range and variance, respectively 0.030 and 0.005, are indiscernible from a κ of 0.

The average κ-value for the A-clade of the A(BC)-phylogeny is 0.70 (see figure 2.20). This value is only significantly different from $\kappa = 1$ in 24 out of 50 trees (p-values ranging from $p < 0.05$ to $p < 0.001$). It provides a better fit than a standard κ-value in 46 trees. This value of κ is relatively high, reaching the default κ-value of 1, which indicates gradual character evolution. The κ-values of the other two clades are lower: 0.06 and 0.40 for clades B and C, respectively. For these clades, the calculated value is indiscernible from a κ-value of 0.

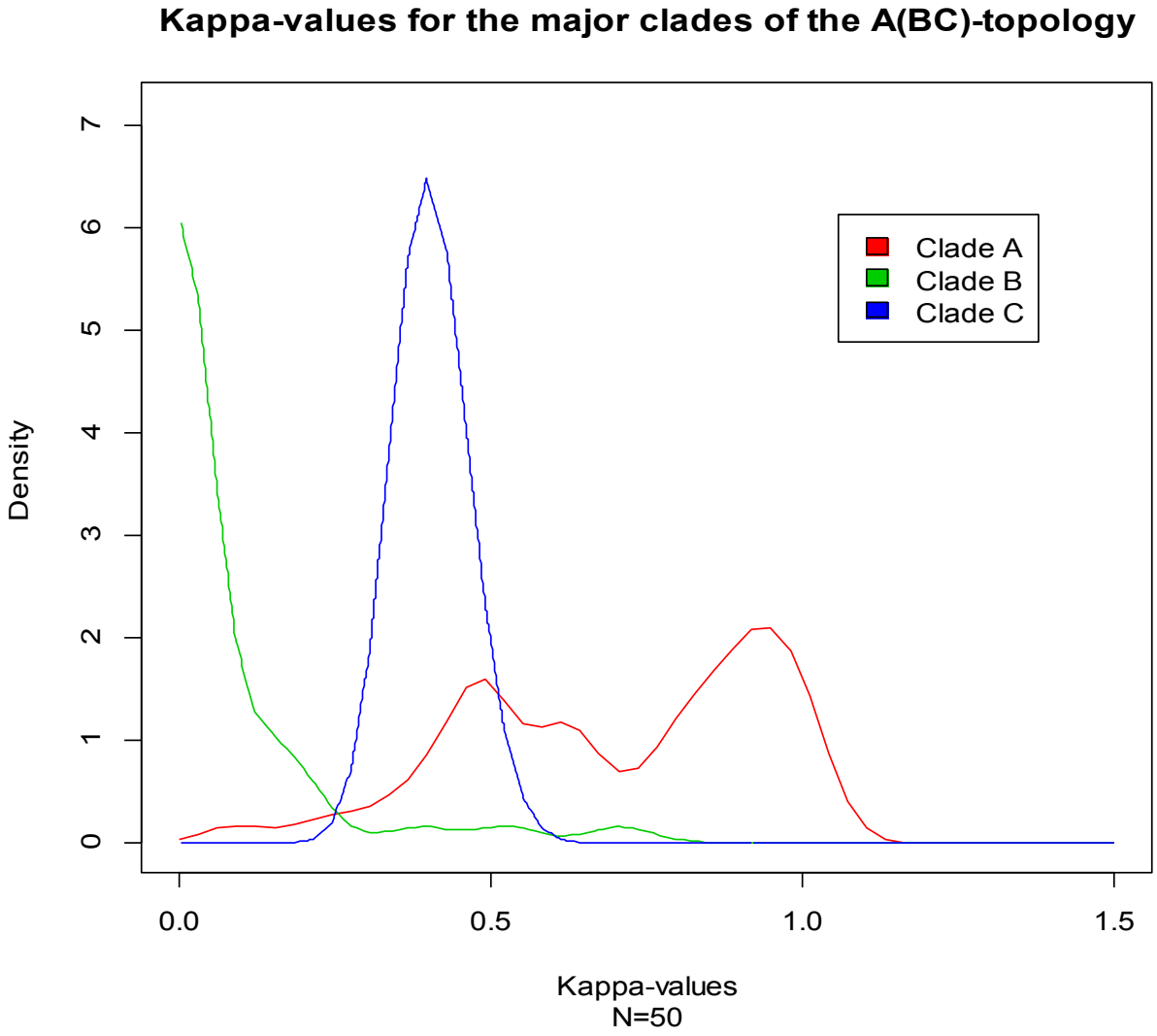

Figure 2.20. κ-values for the major clades of 50 trees of the A(BC)-topology. For graphical reasons, the x-axis (marking the κ-value) ranges from 0 to 1.5, not to the maximum κ-value of 3.

The same pattern can be seen in the (AB)C-topology (see figure 2.21). The average κ-value of clade A is 0.56, and for clades B and C the average value is 0.06. In these last clades, the calculated κ-

value is not any different from a κ-value of 0. In the A-clade, the calculated κ-value is signficantly different from κ = 0, 1, or 3, and does provide a better model fit than one of these standard values.

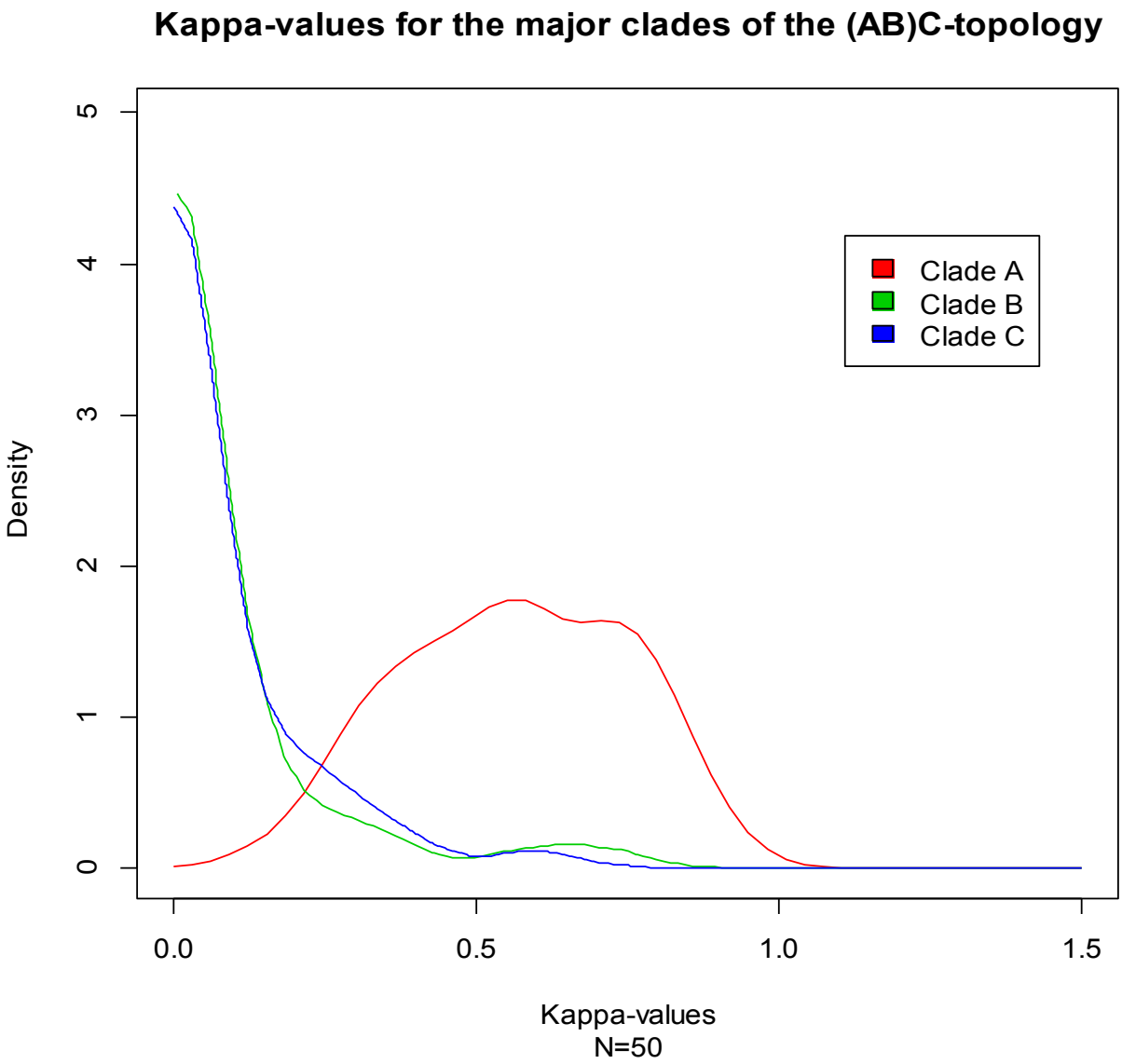

Figure 2.21. κ-values for the major clades of 50 trees of the A(BC)-topology. For graphical reasons, the x-axis (marking the κ-value) ranges from 0 to 1.5, not to the maximum κ-value of 3.

Pagel's δ

Pagel's δ provides a measure for the tempo of character evolution; the tempo of evolution can have decreased (δ = (almost) 0), increased (δ = 3), or remained stable (δ = 1, default) over time. Once again, I fitted a δ-model of evolution to 50 randomly-chosen trees from both topologies, using median spur lengths.

The results are not exactly similar to the results for the other analyses. For both topologies, and also for each major clade, no matter what measure of nectar spur length was used, the calculated δ-value is 2.99 for every tree. This would mean that the tempo of character evolution has strongly increased over time, indicating later species adaptation. Remarkably, for each measure of spur length, δ = 2.99

is significantly different ($p < 0.001$) to $\delta = 3$, and also provides a better model-fit than any of the standard values of δ, in about half of the trees.

Tests with randomly generated trees and data showed that a δ-value of 2.99 is always the outcome when using an ultrametric tree. Apparently, fitting Pagel's δ with "fitContinuous" only works with non-ultrametric trees, even though in its original description, Pagel (1999) never states that δ can not be calculated with ultrametric trees. Nevertheless, the results concerning Pagel's δ in this thesis should be ignored.

Overview

			Topology	
Tree	**Transformation**	**Spur length measure**	**A(BC)**	**(AB)C**
Complete tree	Pagel's λ	median	0.69	0.67
		minimum	0.73	0.69
		maximum	0.54	0.53
		range	0.19	0.23
		variance	0	0
Clade A		median	0.68	0.65
Clade B			0.57	0.60
Clade C			0.40	0.42
Complete tree	Pagel's κ	median	0.35	0.22
		minimum	0.47	0.40
		maximum	0.22	0.10
		range	0.10	0
		variance	0	0
Clade A		median	0.70	0.56
Clade B			0.06	0.06
Clade C			0.40	0.06

Table 2.5. Average λ and κ-values calculated for both topologies, using different measures of spur length.

Model fit

For both topologies, I used "fitContinuous" from the GEIGER-library to find the model that best fits the evolution of the nectar spur lengths over the phylogenetic tree. Available models in fitContinuous are standard Brownian-motion, Pagel's λ, κ, and δ Transformations, an Ornstein-Uhlenbeck model, an early burst model and a white noise model. Once again, I used different measures of spur length; minimum, maximum and median spur length, as well as the variance and

range. The best model is chosen based on AIC.

A(BC)-topology

"fitContinuous" was run using median spur length and 200 randomly-chosen trees from the A(BC)-topology. For each tree, the λ-model provided the best fit (see figure 2.22). The results for minimum and maximum spur length (run with 50 random trees) are quite similar: the best model for each tree is Pagel's λ. However, for the analysis run with the variance and the range of spur lengths (each run with 50 trees), the results are a different: for 34 trees, the model that best describes the evolution of nectar spur range is Pagel's λ, for 9 trees it is the Ornstein-Uhlenbeck model, and for the 7 remaining trees the white noise model is the best model. When the measure of spur length is the variance, for 49 trees the best model is the white noise model, and for the remaining tree it is the Ornstein-Uhlenbeck model. These last results are not a big surprise: as we have seen in the previous section, the λ-value calculated for the variance of nectar spur length was indiscernible from 0. This means that there is no phylogenetic signal, indicating that the topology of the tree does not convey any information about the evolution of the variance. In other words, the character values for each species could have been drawn randomly from a normal distribution. And this is exactly what happens under the white model of evolution: the white model and Pagel's λ model with a λ of 0 are the same.

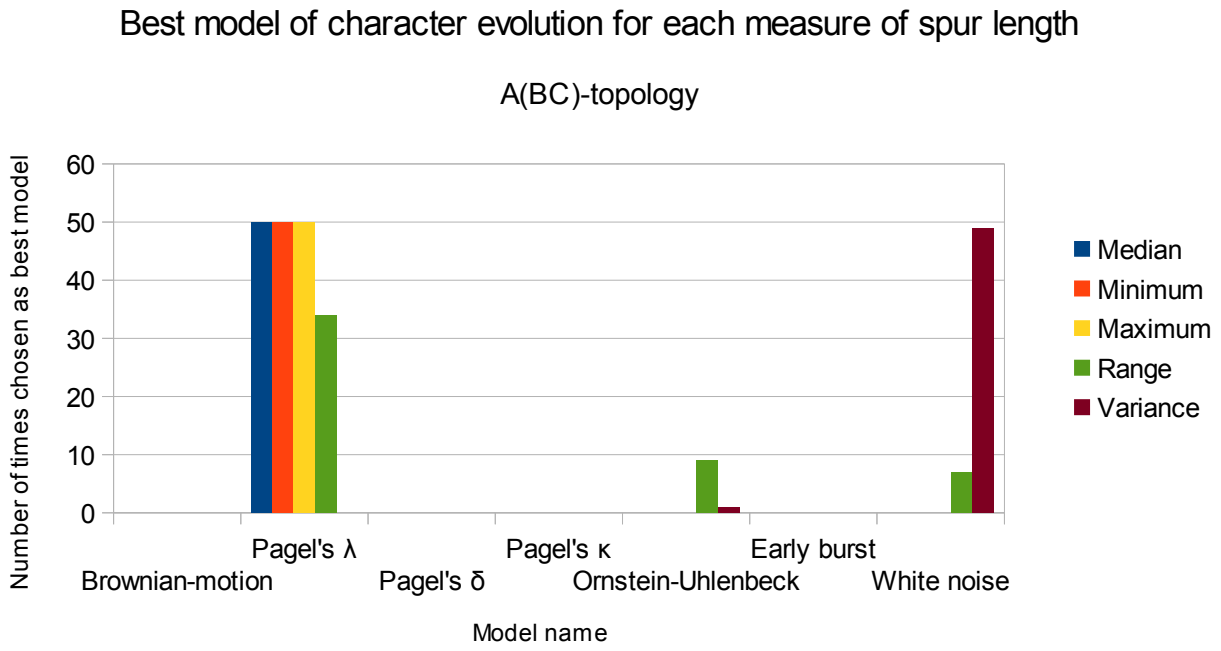

Figure 2.22. The models that best describe the evolution of different measures of nectar spur length over the A(BC)-topology. For each measure of nectar spur length, a model was run over 50 randomly-chosen trees (except for median spur length, which was run over 200 trees). Bars represent the number of times a model was chosen as the best model by AIC. For graphical reasons, the maximum height of the bar belonging to median spur lengths is 50 rather than 200.

As with Pagel's Transformations, "fitContinuous" was also used with just a subset of the whole phylogeny; it was run with as input just one of the major clades A, B or C, of 50 trees, and with the median of spur length. For the A-clade of 43 trees, the model that best describes the evolution of spur length is Pagel's λ, for 1 tree it is the standard random walk-model, and for 6 trees an Ornstein-Uhlenbeck model provides the best fit to the data. Results are more mixed for the B-clade, with 26

trees being best described by Pagel's λ, 17 trees by Pagel's κ, 4 trees by the early burst-model, and 1 tree by simple Brownian-motion, Pagel's δ and Ornstein-Uhlenbeck. The C-clade of 37 trees is best described by the λ-model, whereas an Ornstein-Uhlenbeck-model provides the best fit for the remaining 13 trees. See figure 2.23.

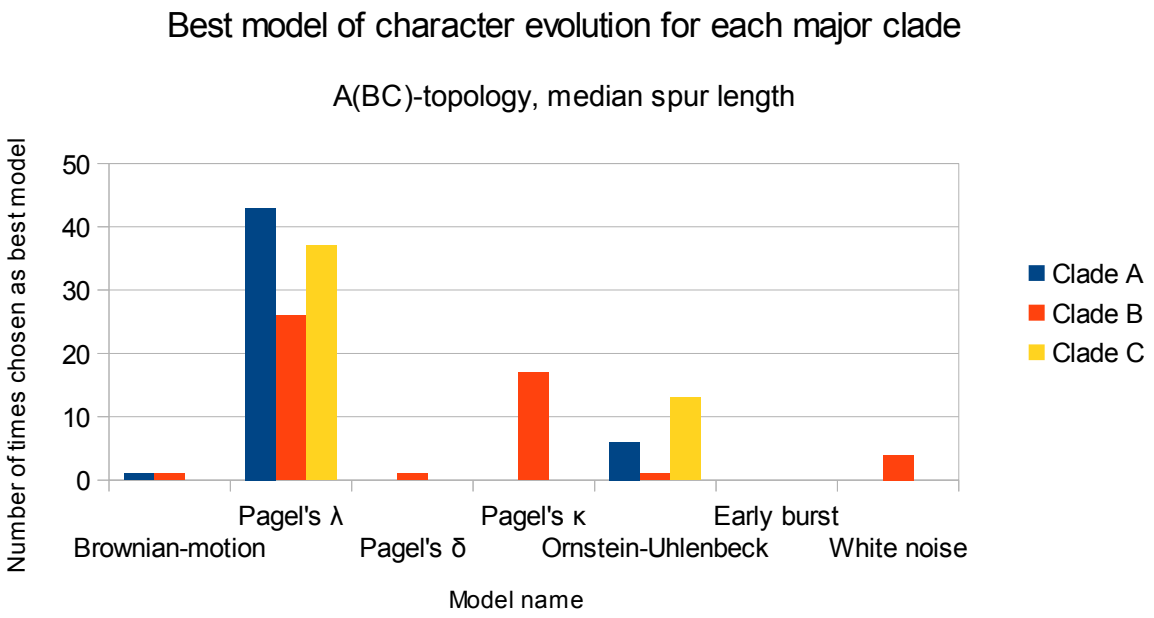

Figure 2.23. The models that best describe the evolution of median spur length for the tree major clades of the A(BC)-topology. For each clade, a model was run over 50 randomly-chosen trees. Bars represent the number of times a model was chosen as the best model by AIC.

(AB)C-topology

Similarly to the analyses performed on the A(BC)-topology, "fitContinuous" was run over 200 trees from the (AB)C-topology using median spur length, and over 50 trees using minimum, maximum, range and variance of spur length. Again, Pagel's λ-model best describes the evolution of the median and maximum of spur length for each tree, just as it does for most of the trees with the minimum (exception: 1 Ornstein-Uhlenbeck-model) and range (exceptions: 4 Ornstein-Uhlenbeck-models). The evolution of the variance of nectar spur length is once again best described by the white noise-model, with the exception of 1 tree, where an Ornstein-Uhlenbeck-model provides a better fit. See figure 2.24.

For 45 trees containing only the A-clade or the C-clade, the best model fit is the λ-model, and for the remaining 5 trees, the Ornstein-Uhlenbeck-model best describes the evolution of median spur length (see figure 2.25). For the B-clade, the results are more mixed: the best model for 17 trees is Pagel's κ, the evolution on 16 other trees is best described the white noise-model, for 14 trees it is Pagel's λ, and for one tree each the best model fit is respectively Brownian-motion, Pagel's δ, and Ornstein-Uhlenbeck.

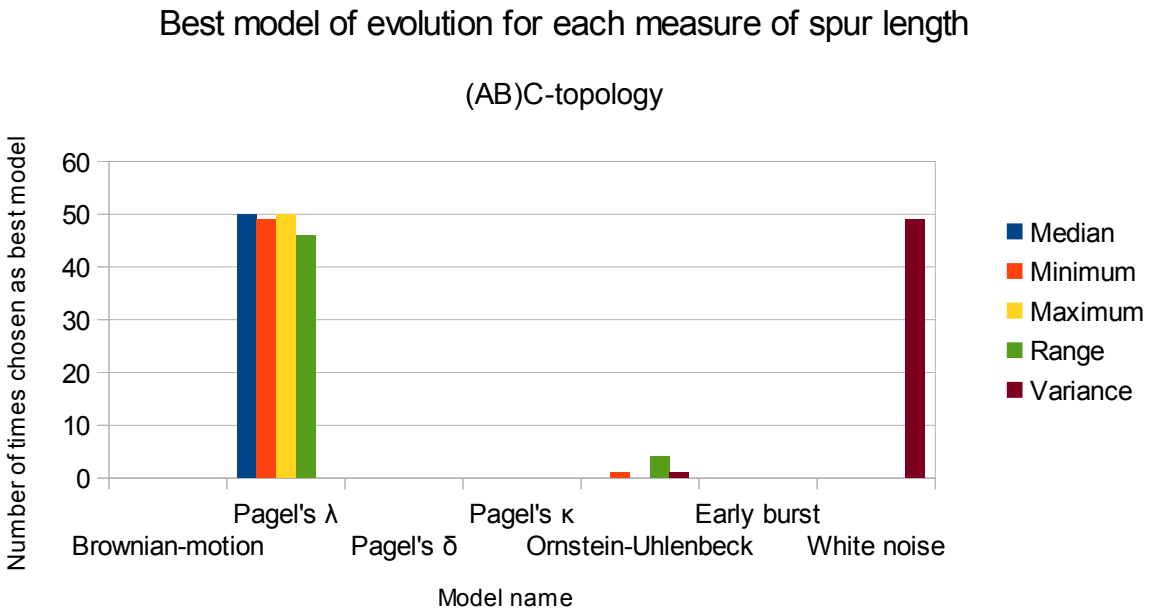

Figure 2.24. The models that best describe the evolution of different measures of nectar spur length over the (AB)C-topology. For each measure of nectar spur length, a model was run over 50 randomly-chosen trees (except for median spur length, which was run over 200 trees). Bars represent the number of times a model was chosen as the best model by AIC. For graphical reasons, the maximum height of the bar belonging to median spur lengths is 50 rather than 200.

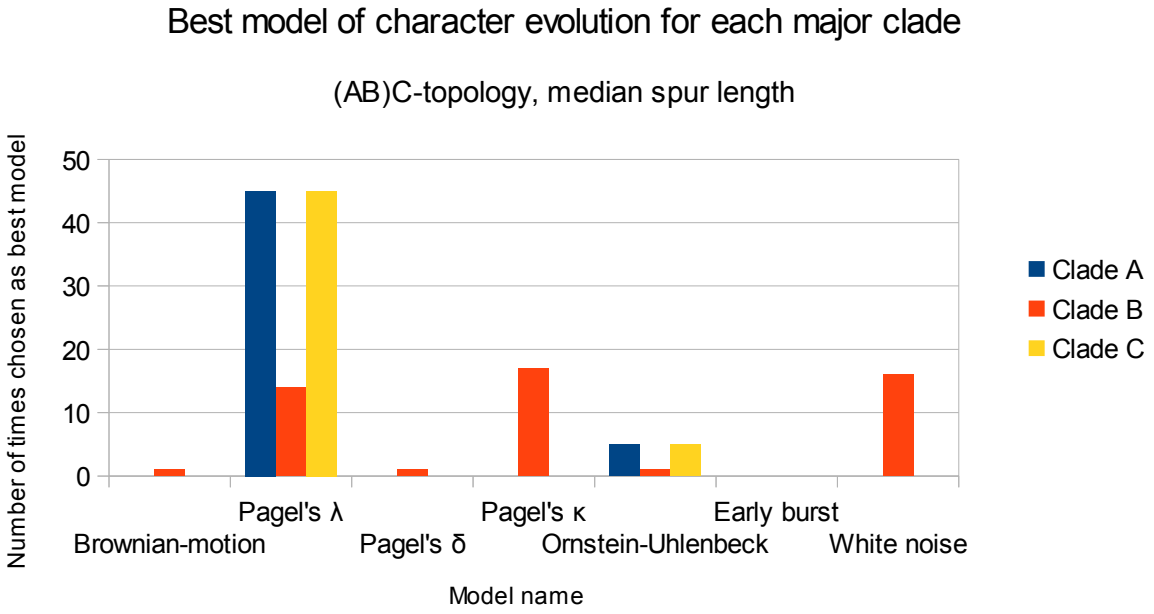

Figure 2.25. The models that best describe the evolution of median spur length for the tree major clades of the A(BC)-topology. For each clade, a model was run over 50 randomly-chosen trees. Bars represent the number of times a model was chosen as the best model by AIC.

		Topology	
Tree	**Measure of spur length**	**A(BC)**	**(AB)C**
Complete tree	Median	Pagel's λ (200/200)	Pagel's λ (200/200)
	Minimum	Pagel's λ (50/50)	Pagel's λ (49/50)
	Maximum	Pagel's λ (50/50)	Pagel's λ (50/50)
	Range	Pagel's λ (34/50)	Pagel's λ (46/50)
	Variance	White noise (49/50)	White noise (49/50)
Clade A	Median	Pagel's λ (43/50)	Pagel's λ (45/50)
Clade B	Median	Pagel's λ (26/50)	Pagel's κ (45/50)
Clade C	Median	Pagel's λ (37/50)	Pagel's λ (45/50)

Table 2.6. Table showing the best model of character evolution for different measures of spur length, topologies, and clades. The model that was most often chosen as the best model is shown. "Pagel's λ (50/50)" means that Pagel's λ-model was chosen as the best model for 50 trees out of a total of 50 trees.

Ancestral spur length estimations

Ancestral spur lengths were estimated for the nodes leading up to the six major clades (C1, C2, B, A1, A2POC - sections *Polyactium*, *Otidia*, and *Cortusina* - and A2HLM - sections *Hoarea*, *Ligularia*, and *Magnistipulaceae*) as well as for deeper nodes, leading up to two or more of these clades. This was done for 200 λ-transformed trees of both topologies, using minimum, median and maximum nectar spur lengths, and three optimization methods: Maximum Likelihood (ML), Restricted Maximum Likelihood (REML), and Squared-Change Parsimony (SCP). The results for the A(BC)-topology can be seen in table 2.7 and figure 2.26, and the results for the (AB)C-topology are shown in table 2.8 and figure 2.27.

	Maximum Likelihood			Restricted Maximum Likelihood			Squared-Change Parsimony		
Clade	**Min**	**Median**	**Max**	**Min**	**Median**	**Max**	**Min**	**Median**	**Max**
A2HLM	1.745	2.263	3.092	1.776	2.431	3.153	2.131	2.851	3.612
A2POC	1.241	1.860	3.761	1.274	1.824	2.499	1.065	1.559	2.088
A1	0.909	1.532	3.488	0.935	1.403	2.037	0.720	1.087	1.514
B	0.823	1.315	4.630	0.849	1.290	1.878	0.593	0.980	1.381
C2	1.505	2.697	2.829	1.517	2.129	2.863	1.270	1.912	2.482
C1	1.188	2.058	2.754	1.180	1.833	2.796	1.115	1.909	2.762
A2HLM + A2POC	1.241	1.860	3.761	1.274	1.824	2.499	1.065	1.559	2.088
A2 + A1	1.241	1.860	3.761	1.274	1.824	2.499	1.065	1.559	2.082
C1 + C2	1.280	2.288	3.039	1.287	1.910	2.758	1.100	1.754	2.409
B + C	1.089	1.784	4.344	1.287	1.639	2.307	0.903	1.424	1.959
A + B + C	1.141	1.820	4.256	1.172	1.704	2.375	0.976	1.489	2.025

Table 2.7. Ancestral nectar spur lengths (in centimetres) reconstructed for nodes leading up to the major clades of the A(BC)-topology. Three measures of spur length and three reconstruction methods were used. In the upper half of the table spur lengths of nodes leading up to individual clades are shown, in the lower half of the table spur length reconstructions for the most recent common ancestor of two or more clades are depicted (e.g., "B+C" means the node leading up to the B- and C-clades). "A+B+C" is the reconstruction for the most basal node in the phylogeny.

	Maximum Likelihood			Restricted Maximum Likelihood			Squared-Change Parsimony		
Clade	**Min**	**Median**	**Max**	**Min**	**Median**	**Max**	**Min**	**Median**	**Max**
A2HLM	1.787	2.501	3.196	1.795	2.446	3.148	1.889	2.577	3.277
A2POC	1.276	2.061	2.664	1.257	1.797	2.449	1.023	1.514	2.017
A1	0.971	1.583	2.311	0.978	1.447	2.069	0.741	1.108	1.522
B	0.969	1.949	2.364	0.858	1.304	1.882	0.641	1.046	1.440
C2	1.664	2.206	2.433	1.493	2.111	2.819	1.317	2.032	2.623
C1	1.341	2.126	2.583	1.178	1.821	2.735	1.141	1.950	2.809
A2HLM + A2POC	1.266	2.050	2.652	1.247	1.786	2.437	1.018	1.511	2.014
A2 + A1	1.263	2.049	2.649	1.248	1.783	2.433	0.991	1.474	1.968
A + B	1.166	2.152	2.519	1.063	1.569	2.208	0.889	1.381	1.870
C1 + C2	1.438	2.316	2.532	1.265	1.877	2.677	1.160	1.868	2.544
A + B + C	1.319	2.332	2.566	1.165	1.725	2.445	1.026	1.626	2.208

Table 2.8. Ancestral nectar spur lengths (in centimetres) reconstructed for nodes leading up to the major clades of the (AB)C-topology.

Several conclusions can be drawn based on these results.
First of all, the average results from the different reconstruction methods are fairly similar. There are some exceptions (most notable the maximum spur lengths calculated by maximum likelihood for the A(BC)-topology), but the differences between the reconstruction methods hardly ever exceed more than half a centimetre. However, when looking at the individual trees (as can be seen in figures 2.28 and 2.29), the differences between the optimization methods may be quite striking. Only when the nectar spur reconstructions are averaged over all the trees do the differences become smaller.
Secondly, the results for the different clades are quite similar as well. Almost all major clades have a median ancestral spur length between 1 and 2.5 centimetres. The origins of the major clades seem to have had neither extremely short nectar spurs, probably corresponding with pollinators such as bees, nor extensively long nectar spurs, which are pollinated by pollinators with long proboscids.

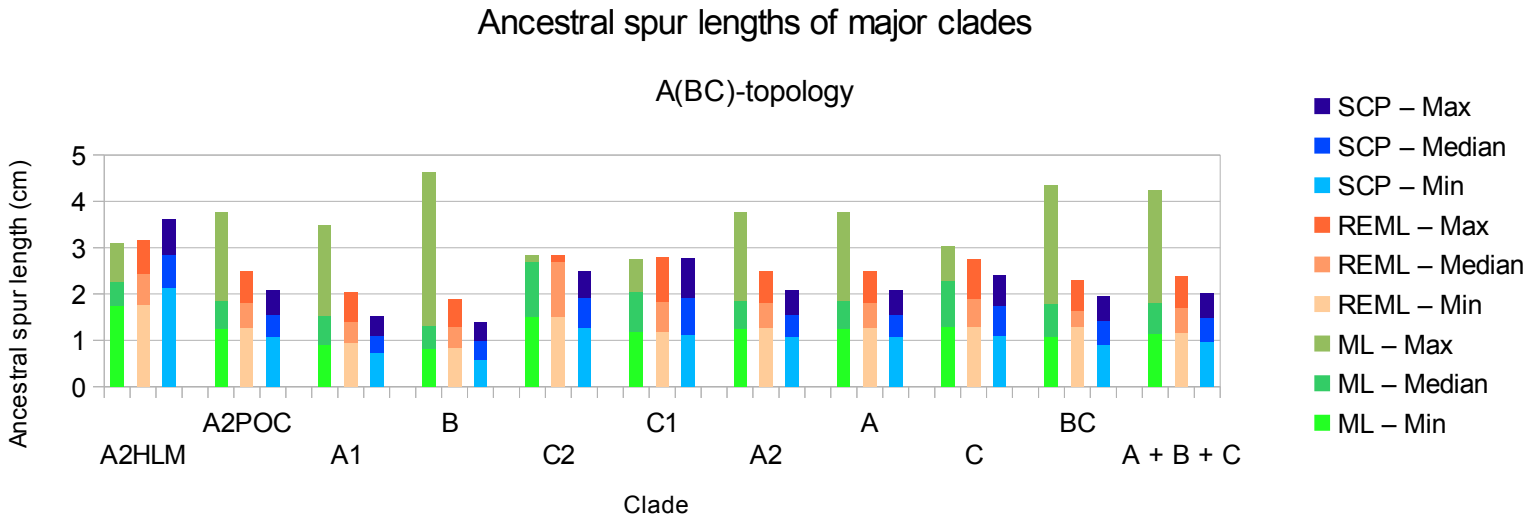

Figure 2.26. Ancestral spur lengths of the nodes leading up to the major clades of the A(BC)-topology. Green bars represent minimum, median and maximum spur lengths reconstructed using Maximum Likelihood, red bars depict spur lengths reconstructed by Restricted Maximum Likelihood, and blue bars represent spur lengths reconstructed by Squared-Change Parsimony.

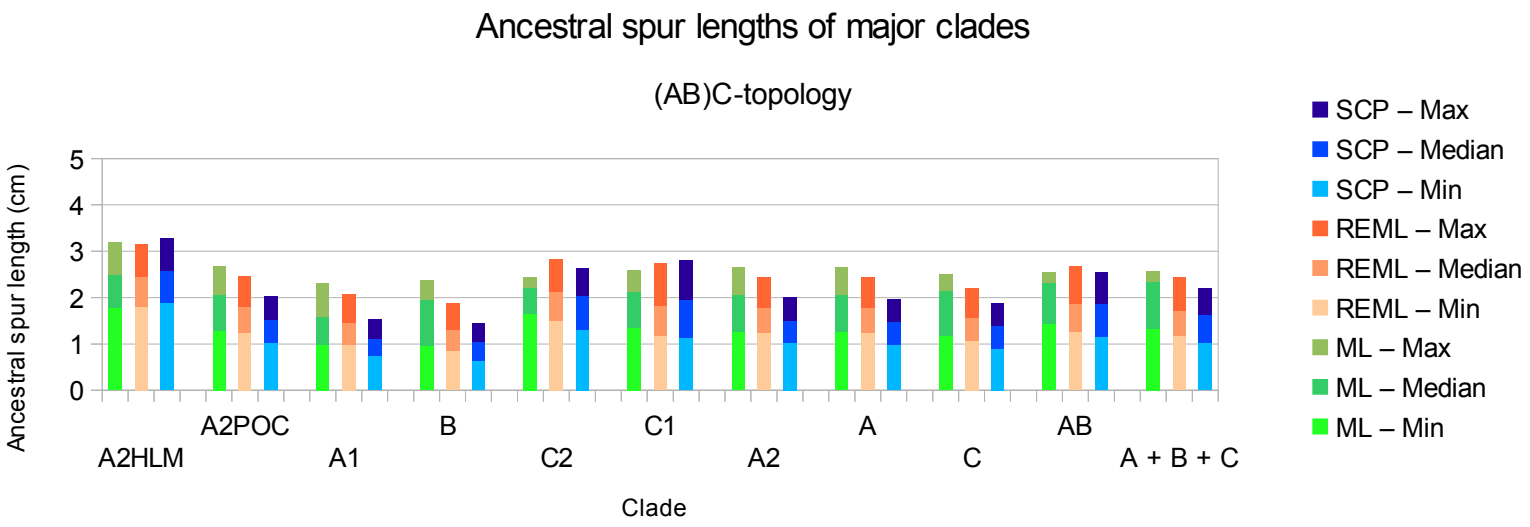

Figure 2.27. Ancestral spur lengths of the nodes leading up to the major clades of the (AB)C-topology.

Thirdly, the median nectar spur length of the most basal node in the tree, representing the most ancestral *Pelargonium* species, probably had a length of 1.5 to 2.3 centimetres, with a possible maximum value as high as 4.3 centimetres, and a minimum value of 1.0 centimetre. Even though this is still quite a large range, one firm conclusion can be drawn based on these results: the nectar spurs of the ancestral *Pelargonium* were significantly longer than previously calculated (Bakker *et al.* 2005). This previous reconstruction of nectar spur length resulted in an ancestral *Pelargonium* with a spur length of 0.5 centimetres, whereas even the smallest reconstructed nectar spur in this thesis is at least twice as long.
A final result is that based on these reconstructions, it is hard to distinguish a trend towards larger (or shorter spurs). The most basal node in the tree does not have much smaller nectar spurs than most other, more distal nodes. In fact, for some reconstructions, the reconstructed spurs are even slightly longer than in the basal node than in the node leading up to clades B and A1.

Phylogenetic uncertainty

To take into account phylogenetic uncertainty, ancestral states were reconstructed for 400 trees in total (200 trees per topology), after which the results were averaged. As a measure of the differences between the three reconstruction methods the sum of squared differences between two methods were calculated for each tree. The results can be seen in figures 2.28 and 2.29.

Several conclusions can be drawn from these results.
First of all, for all trees, the differences between the squared-change parsimony analysis and the restricted maximum likelihood analysis are relatively small. The biggest differences are between maximum likelihood on the one hand and parsimony or restricted maximum likelihood on the other hand.
Secondly, the results vary widely per tree. For some trees, all analyses yield fairly similar results, while for some other trees, the differences between the different methods are quite large.
Finally, in general the results for the trees of the (AB)C-topology are more similar than the results from the A(BC)-topology (note the difference in scale of the y-axis).

It can be concluded that, as the results vary greatly per reconstruction method and per tree, it is recommendable to use a variety of methods, and a range of trees.

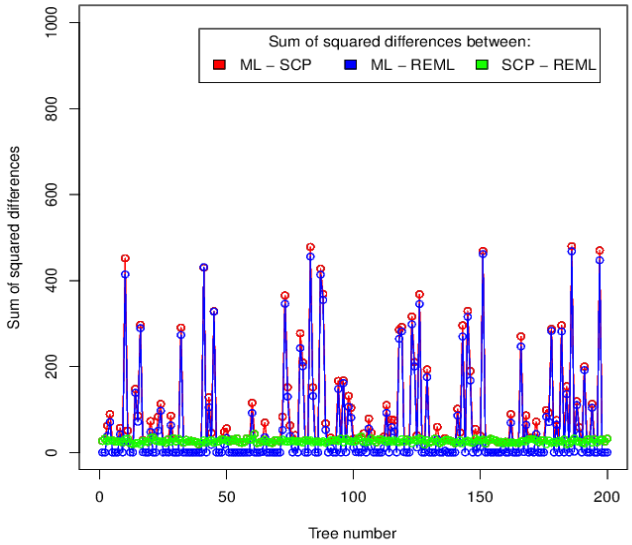

Figure 2.28. Sum of squared differences between different reconstruction methods for the (AB)C-topology. Abbreviations: ML - Maximum Likelihood, REML - Restricted Maximum Likelihood, SCP - Squared-Change Parsimony.

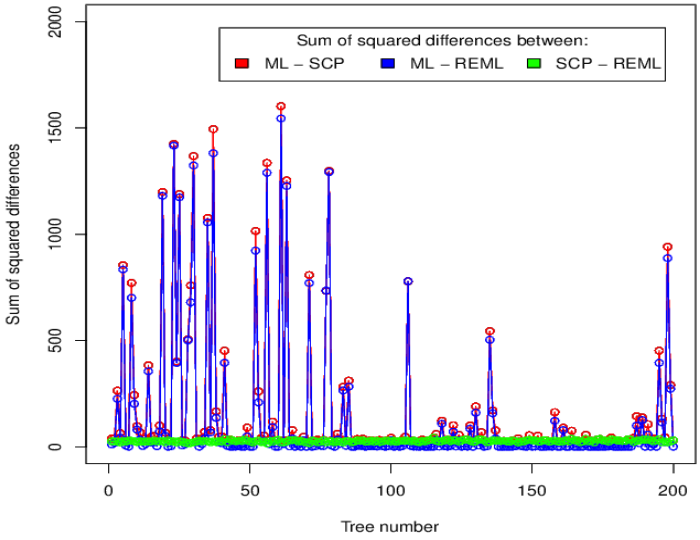

Figure 2.29. Sum of squared differences between different reconstruction methods for the A(BC)-topology. (Please note the different scale of the y-axis compared to figure 2.28.)

A final note on the suitability of the used methods: four times I reconstructed the nectar spur length of the most ancestral node of 200 trees of the (AB)C-topology with randomized spur lengths, to test whether this affects the outcome of the ancestral reconstructions. As can be seen in table 2.9, the results varied per analysis, indicating that the reconstruction of the most ancestral node in the tree is not simply the average of all species values. Instead, the distribution of nectar spur values of the species in the phylogeny has an effect on the outcome of the reconstruction analysis. I regard this as evidence that the methods used for reconstructing ancestral nectar spur lengths provide valuable results.

		Randomized values			
Reconstruction method	**Original values**	**1**	**2**	**3**	**4**
SCP	1.63	2.16	1.78	2.11	2.25
ML	2.33	1.19	1.89	1.83	2.59
REML	1.73	1.99	1.95	1.94	2.25

Table 2.9. Reconstructed nectar spur lengths (in centimetres) of the most ancestral node of 200 trees of the (AB)C-topology per used method. Both results from the previous reconstruction (see above) and four randomized reconstructions are shown.

Stochastic character mapping

SIMMAP was run four times, to check whether the amount of mutational maps was sufficient. The results of the four runs were identical, indicating the four runs have converged on a single outcome.

The average number of changes between states varied between 1991 and 2139, depending on the run. It is important to note that the only changes that occurred were between successive states; a

change from state 1 to state 3, for example, was never mapped. Nevertheless, for each run the average number of changes towards a higher state was higher than the number of changes towards a lower state. This could indicate a trend towards longer nectar spurs. However, the differences between changes to higher and lower states were small (varying between 16 and 18), so the trend was not very clear. Furthermore, a different definition of the character states may cause different results, as changes of nectar spur length within one state (e.g., a change from a spur length of 1.5 centimetres to a nectar spur of 2.2 centimetres) are not identified by SIMMAP.

SIMMAP calculates the average dwelling time per character state (see figure 2.30). As can be seen, most time is dwelt in states 1, 2, and 3, corresponding with a median nectar spur length between 0 and 3.6 centimetres.

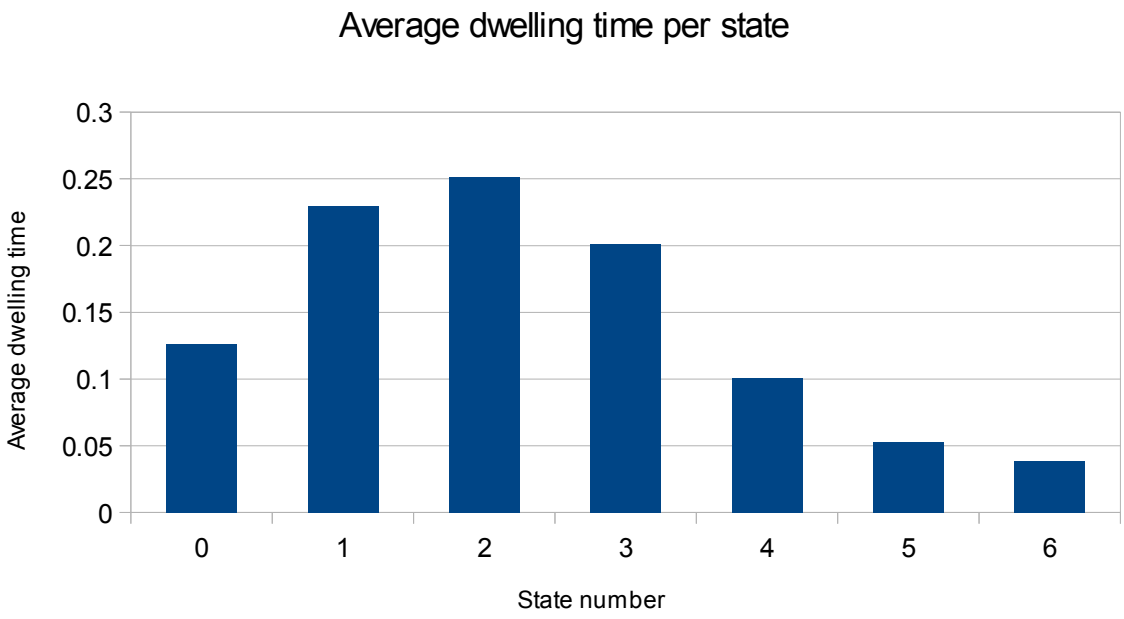

Figure 2.30. Average dwelling time per character state. (Definition of character states: median nectar spur length in centimetres = x. State 0: x = 0. State 1: $0 < x \leq 1$, state 2: $1 < x \leq 2.3$, state 3: $2.3 < x \leq 3.6$, state 4: $3.6 < x \leq 4.9$, state 5: $4.9 < x \leq 6.2$, state 6: $6.2 > x$.)

SIMMAP also calculates the probability of being in a specific character state for each node in the tree. Similar to the results from the analyses in R, the probabilities are shown for the main clades (see table 2.10 and figure 2.31).

Some results are quite remarkable. First of all, the ancestral spur length of the A2-clades has a high probability of being between 1 and 3.6 centimetres, whereas for clades A1 and B an ancestral spur length between 0 and 1 centimetre is more likely. The results for the C-clades are less pronounced, with the highest probability for an ancestral spur length between 1 and 2.3 centimetres.
Secondly, the results for the most basal node in the phylogeny, the hypothetical ancestral *Pelargonium*, are not very straightforward. There is a probability of approximately 68% that the ancestral spur length was between 0 and 2.3 centimetres, with equal probabilities for each state. Longer ancestral nectar spur lengths seem unlikely. Based on these results, it is very difficult to conclude what the ancestral spur length of *Pelargonium*, and therefore the ancestral pollinator, may have been.
Finally, table 2.10 shows that some clades are not present in all of the 201 trees. The C-clades tend to be present in (nearly) any tree, but the A-clades are only present in approximately 75% of all

trees. This indicates that the monophyly of these clades can still be disputed.

		Probability of being in						
Clade	**Present in number of trees**	**State 0**	**State 1**	**State 2**	**State 3**	**State 4**	**State 5**	**State 6**
A2POC	173	0.125404	0.15945	0.466301	0.176903	0.036067	0.020154	0.01572
A2HLM	143	0.044418	0.04955	0.088458	0.547554	0.184384	0.051404	0.034231
A1	162	0.26202	0.39772	0.202899	0.056043	0.032504	0.025722	0.023092
B	163	0.368514	0.327312	0.158013	0.058811	0.035175	0.027557	0.024619
C1	201	0.155226	0.21409	0.30845	0.14455	0.0664	0.056713	0.054571
C2	199	0.115834	0.157406	0.309264	0.213853	0.081686	0.062293	0.059664
A2 + A1	163	0.174601	0.290191	0.324654	0.129882	0.037258	0.023504	0.019911
A2HLM + A2POC	163	0.078677	0.093471	0.239748	0.444769	0.08987	0.030641	0.022823
A + B	198	0.257109	0.292649	0.211997	0.08942	0.05523	0.048016	0.04558
C1 + C2	199	0.164733	0.225936	0.295326	0.141817	0.065186	0.054669	0.052334
A+B+C	201	0.200825	0.242442	0.236722	0.133547	0.073293	0.05827	0.054901

Table 2.10. Ancestral state probabilities for each major clade as calculated by SIMMAP. For each clade, the probability of each state is shown, as well the number of trees in which the clade was present (total number of trees is 201).

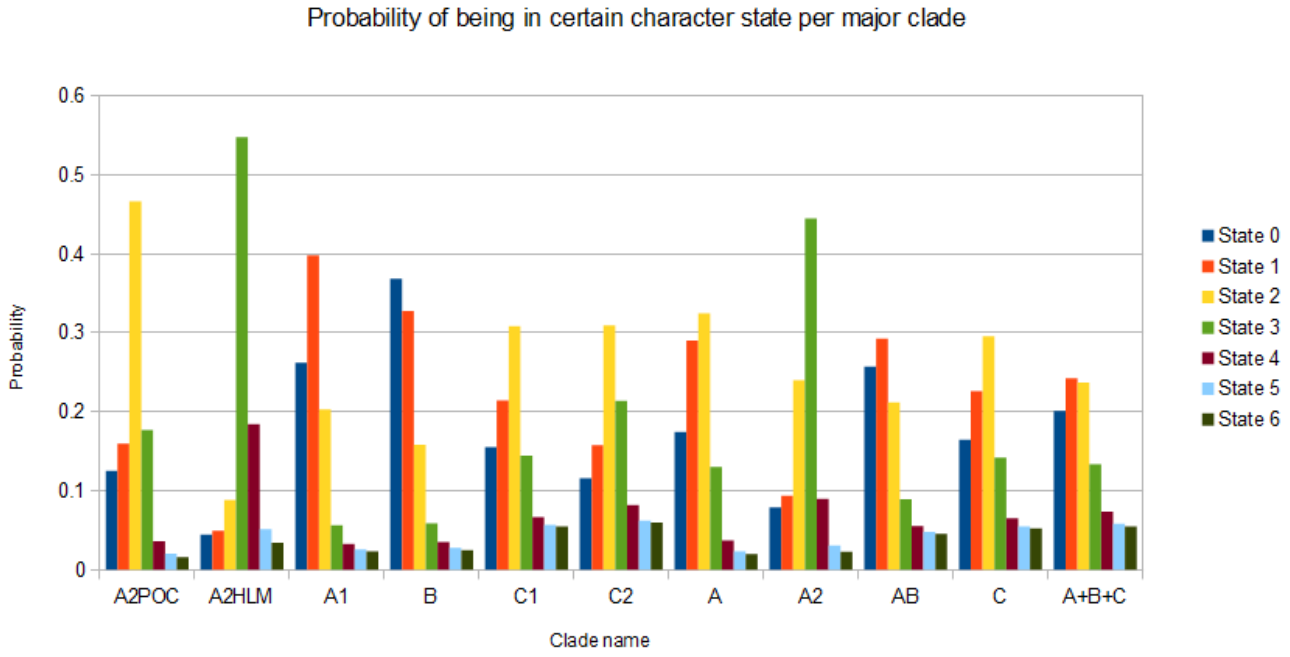

Figure 2.31. Ancestral state probabilities for each major clade. For the exact probabilities, see table 2.10.

Diversification rate shifts

A MEDUSA analysis was performed on 800 trees in total (200 normal trees and 200 skeleton trees per topology). The output of the MEDUSA-analysis consists of several parameters: the number of breakpoints and different tree models, the positions of these breakpoints, and the model parameters. The most important model parameters are the r-value (birth-rate minus death-rate) and the epsilon-value (death-rate divided by birth-rate). In all cases, the epsilon-value was not available (NA). This indicates that no death-rates could be calculated (since there are no fossil species in the trees), which is affirmed by the fact that for all trees the Yule-model was used as tree model, rather than a birth-death model. However, the r-values were available, and since there is no death-rate, the r-values simply depict the birth-rate of the tree model.

As the results of the MEDUSA-analysis vary highly per topology, I will discuss them in different sections.

(AB)C-topology

The results for the skeleton tree of the (AB)C-topology, with every clade represented by one species, are all rather similar: in all 200 trees 1 breakpoint was found. This breakpoint was always located at the foot of the Nanum-clade, which only consists of *Pelargonium nanum* (see figure 2.32). The r-value for this clade was always 0; much lower than the birth-rate for the rest of the tree, which was variable. This indicates that there has been one diversification rate shift in the *Pelargonium*-phylogeny, a shift located in a clade with significantly less speciation.

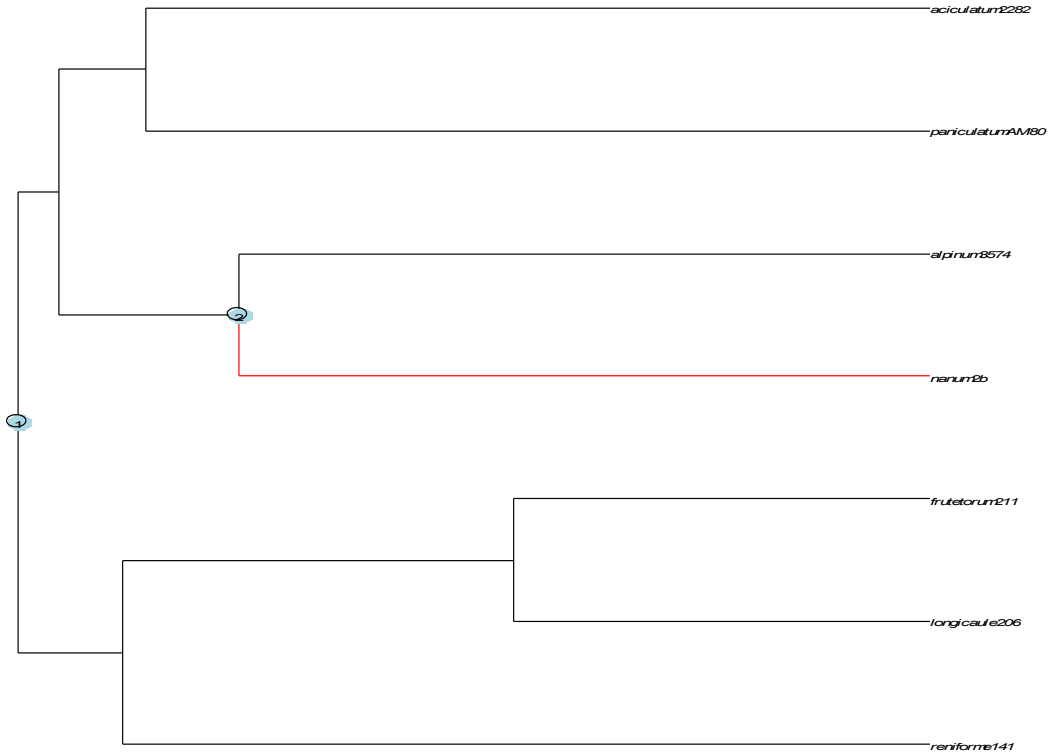

Figure 2.32. Skeleton tree of the A(BC-)topology with diversification rate shifts. Rate shifts are indicated with blue circles. The part of the tree contained by the second rate shift (the red branches) has a birth-rate of 0; significantly lower than the rest of the tree.

However, the results for the full tree, incorporating all species, are highly variable. In 198 out of 200 trees 1 breakpoint was found, indicating the presence of a diversification rate shift. No breakpoints were found in the remaining 2 trees. However, no clear trend could be found considering either the position or the model parameters of this rate shift. In most of the 198 trees, the breakpoint was located in different positions, either incorporating just a few species, or major clades. Similar to the location of the breakpoint, there was much variation in the model parameters as well. For some trees the species or clades contained by the breakpoint had higher birth-rates than the rest of the tree, for other trees they were lower. Based on these results, there is no clear evidence for the presence or absence of a diversification rate shift.

(AB)C-topology

The results for the (AB)C-topology are completely different. For the skeleton tree, 186 of 200 trees had one breakpoint (no evidence for a diversification rate shift was found in the remainder of the trees). In 15 of these 186 trees, the breakpoint was found at the base of the Nanum-clade, with corresponding birth-rates of 0. For the remaining trees, the breakpoint was either located at the base of the A-clades (143 trees, see figure 2.33), the base of the A-clades plus the Nanum-clade (19 trees), or the base of the A2-clade (19 trees). For all these trees, higher birth-rates were found for the A-clades than for the rest of the tree (see figure 2.34).

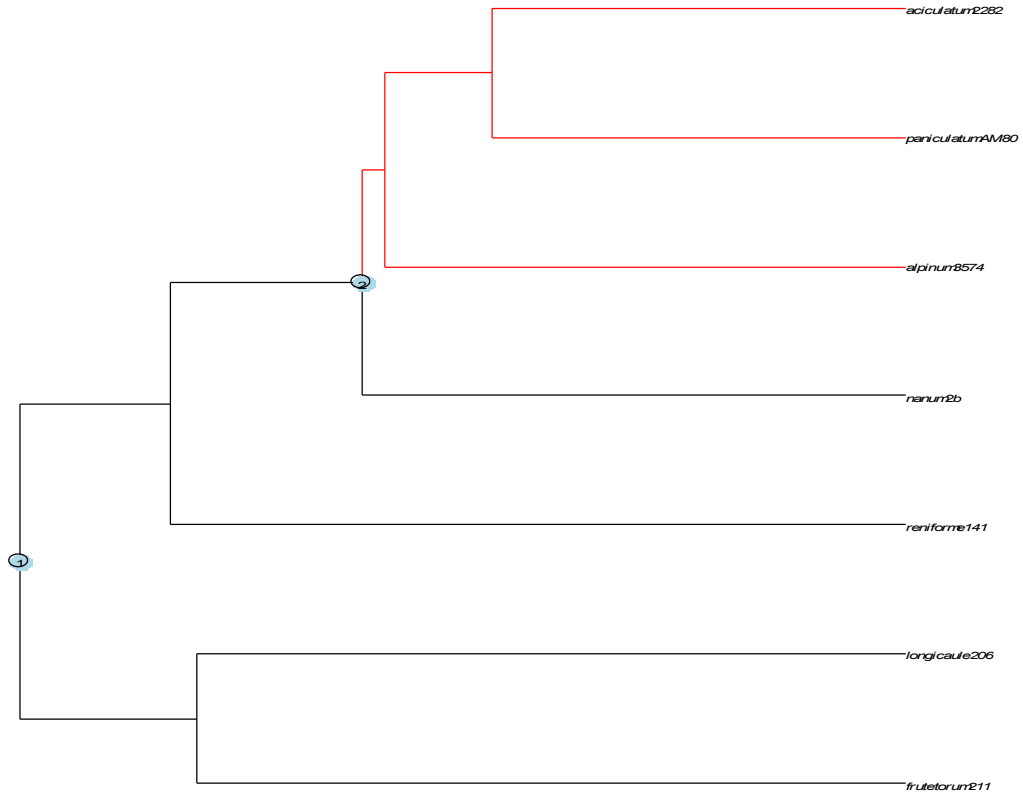

Figure 2.33. Skeleton tree of the (AB)C-topology with diversification rate shifts. Rate shifts are indicated with blue circles. The part of the tree contained by the second rate shift (with the red branches) has higher birth-rates than the rest of the tree.

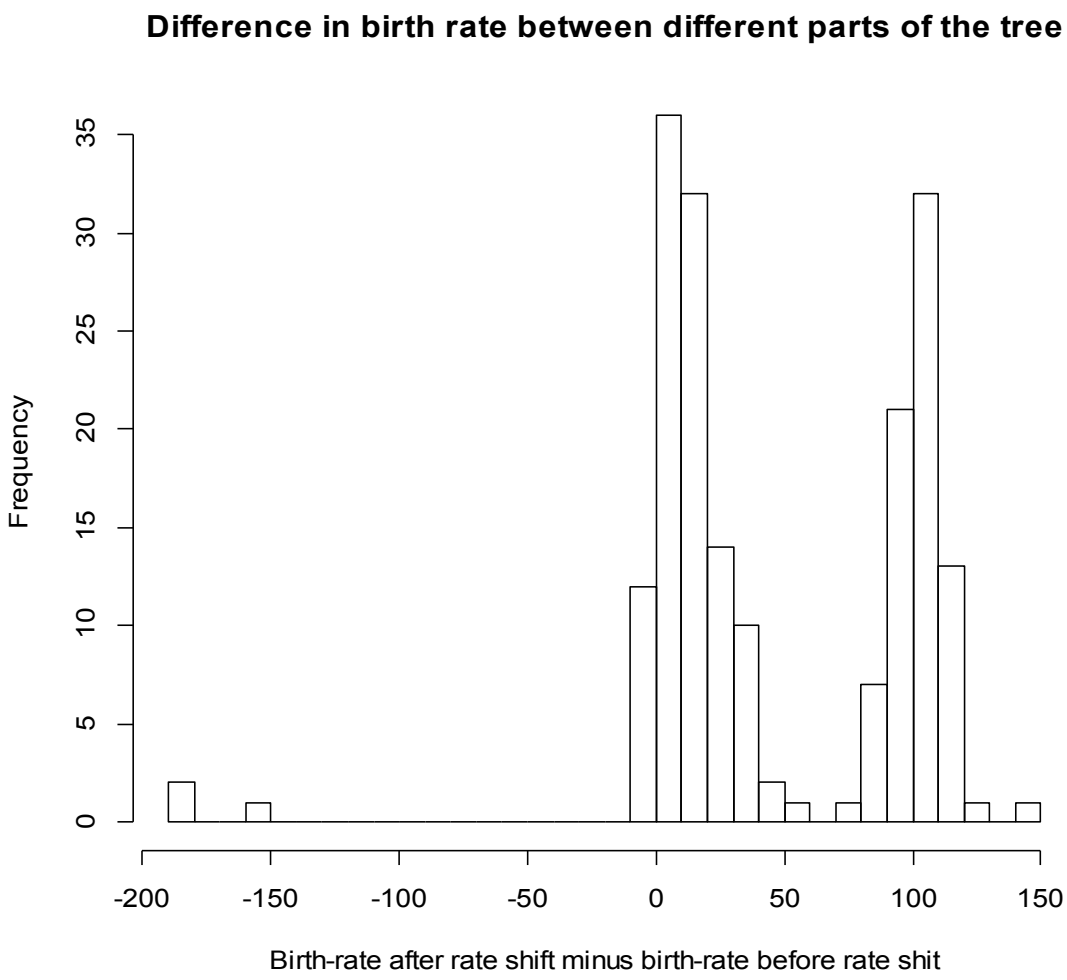

Figure 2.34. Difference in birth-rate before and after the diversification rate shift for the skeleton tree of the (AB)C-topology. As can be seen, birth-rates in clades contained by the diversification rate are almost without exception higher than in the rest of the tree.

When using a full tree rather than a skeleton tree, the location and direction of rate shifts are quite similar, but the amount of shifts found is much lower: in only 48 of 200 trees a diversification rate shift was found. Of these 48 shifts, 33 were found at the base of the A-clades and the Nanum-clade (see figure 2.35). 13 were located at the foot of the A-clades, excluding *P. nanum*. The remaining two shifts were located either at the foot or within the A2-clade. Once again, birth-rates of the clades contained by a rate shift were higher than birth-rates of the rest of the tree (with one exception, see figure 2.36).

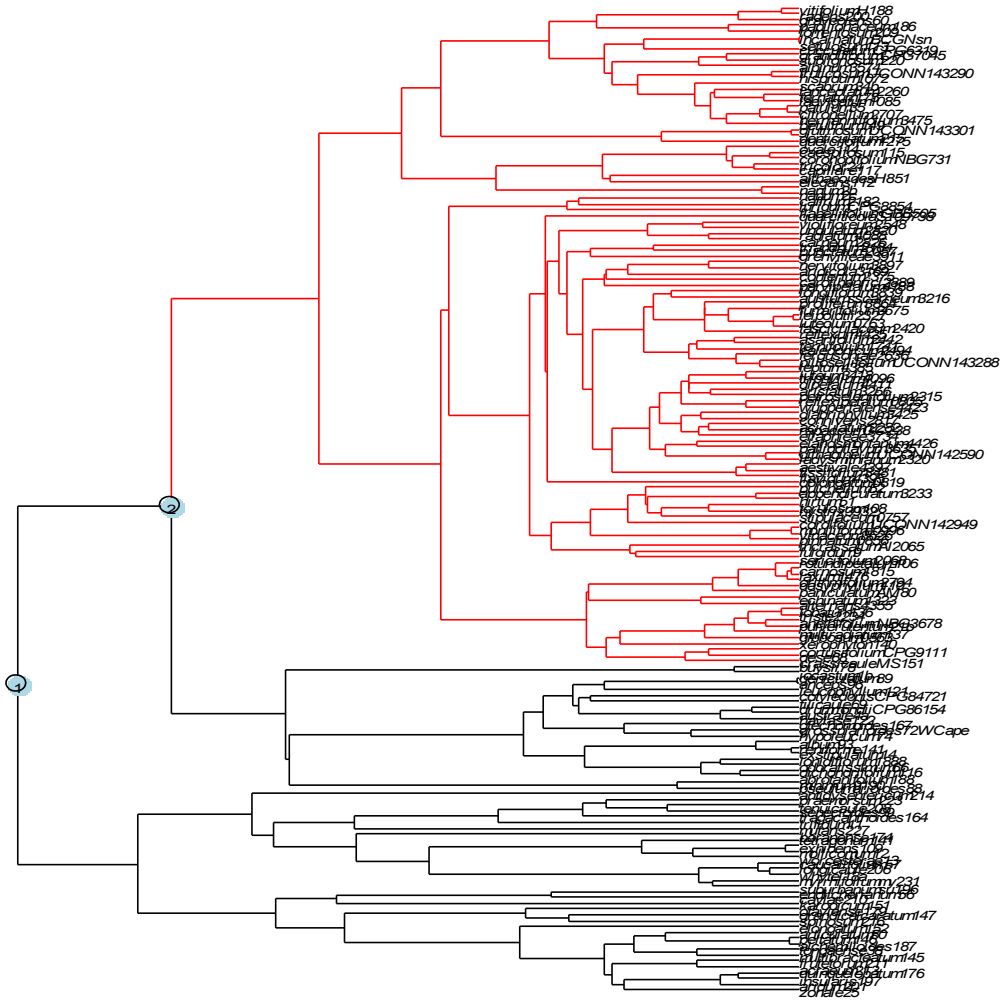

Figure 2.35. Complete tree of the (AB)-topology with diversification rate shifts. Rate shifts are indicated with blue circles. The part of the tree contained by the second rate shift (with the red branches) has higher birth-rates than the rest of the tree.

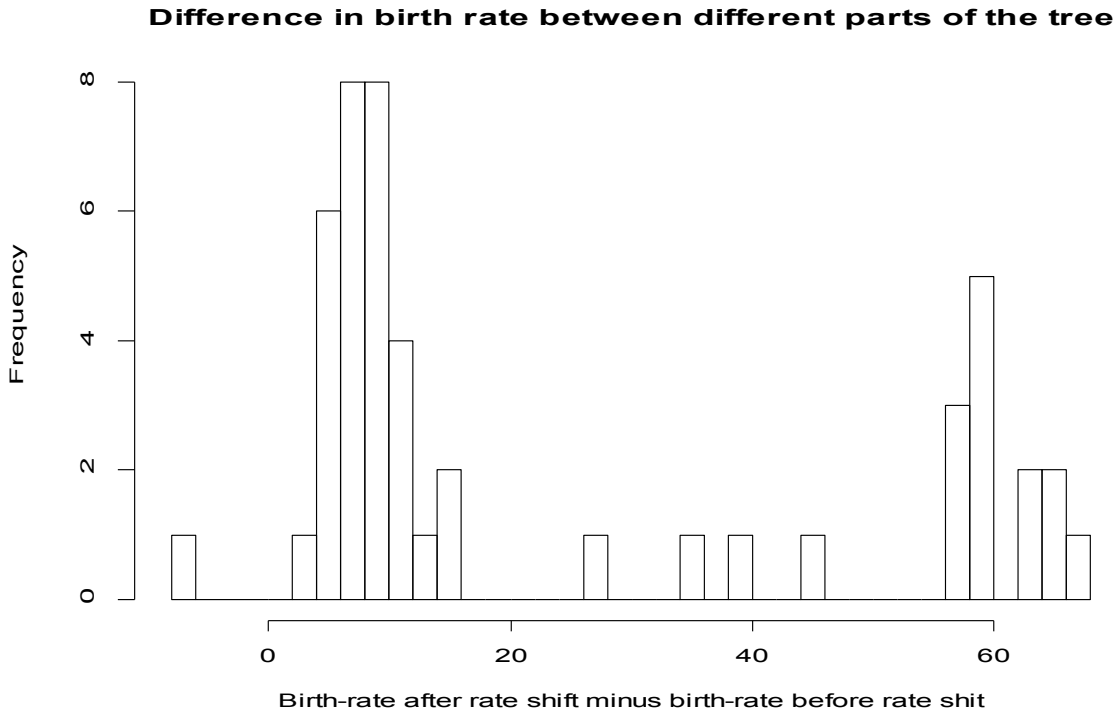

Figure 2.36. Difference in birth-rate before and after the diversification rate shift for the full tree of the (AB)C-topology. Similar to the results from the skeleton tree, birth-rates in clades contained by the diversification rate are higher than in the rest of the tree.

Spur length-dependent speciation rates

The biggest problem of QuaSSE is that its calculations are quite extensive, and take a long time (i.e., per tree the calculations take more than a full day). Therefore, I have not been able to perform a QuaSSE analysis on every available tree. Instead, I had to settle on a compromise between what is computationally possible and what is scientifically desirable.

Using median nectar spur lengths, QuaSSE was run on 20 trees of the A(BC)-phylogeny and 15 of the (AB)C-phylogeny. Additionally, QuaSSE was also run on 5 trees of the A(BC)-phylogeny using minimum, maximum and the range of spur length, and on 15 trees of the (AB)C-phylogeny with the implementation of a diversification rate shift found by MEDUSA.

For every tree and measure of spur length, there were two clear results. First of all, models incorporating an effect of spur length on speciation rates performed significantly better than the null model. This indicates that spur length has a strong effect on speciation rates; an indication of species selection (Rabosky & McCune 2009). Secondly, models incorporating a drift parameter performed significantly better than the models without such a parameter. As the drift parameter was positive for (nearly) every tree and model, this implies that there is a significant evolutionary trend towards longer nectar spurs.

In most cases, several models per tree provided a significant better fit than the null model. Therefore, I chose to incorporate only models that provided a highly significant ($p < 0.001$) better fit than the null model. This way only the best models were taken in account. Of these models, I averaged the model parameters, so for each speciation function (i.e., linear, sigmoid or parabolic) this resulted in one final function, averaged over the results of all trees.

A(BC)-topology

With median spur lengths, QuaSSE was run on 20 trees of the A(BC)-topology. For 11 of these trees, the parabolic function turned out to be the best model. For 4 trees it was the linear model, and for 3 trees the sigmoid model. For 2 trees, there was no model that provided a (highly) significantly better fit ($p < 0.001$) than the null model. As there are three possible speciation functions per tree (i.e., linear, sigmoid, and parabolic), and 20 trees, a total of 60 models were tested. Of these 60 models, 34 were a highly significant improvement over the null model; a fraction of 0.57. The average speciation functions, based on the model fits of these 34 models, can be seen in figure 2.37.

As can be seen, for both the linear and the sigmoid model, longer nectar spurs result in lower speciation rates. This is especially clear in the linear model, where spur longer than 6 centimetres result in a speciation of 0. However, the parabolic model has a slightly different result: again longer spurs tend to have lower speciation rates, but only up to a certain spur length, which is approximately 5 centimetres. Spurs longer than 5 centimetres entail steadily increasing speciation rates. This is an interesting result, but it is important to note that *Pelargonium* species with a median spur length longer than 5 centimetres are quite rate; in fact, only 7 species have nectar spurs that long.

QuaSSE was also run using minimum, maximum and the range of spur length, on 5 trees each. The results can be seen in figure 2.38. When using maximum spur length, 0.80 of all the available models were a highly significant improvement over the null. The parabolic model was chosen 3 times as the best model, the linear model 2 times. For minimum spur length, 0.53 of the models were highly significant. For 4 trees, the parabolic function provided the best fit. For the remaining tree, none of the models was highly significant. For none of the trees, the sigmoid model was a highly significant improvement over the null model.

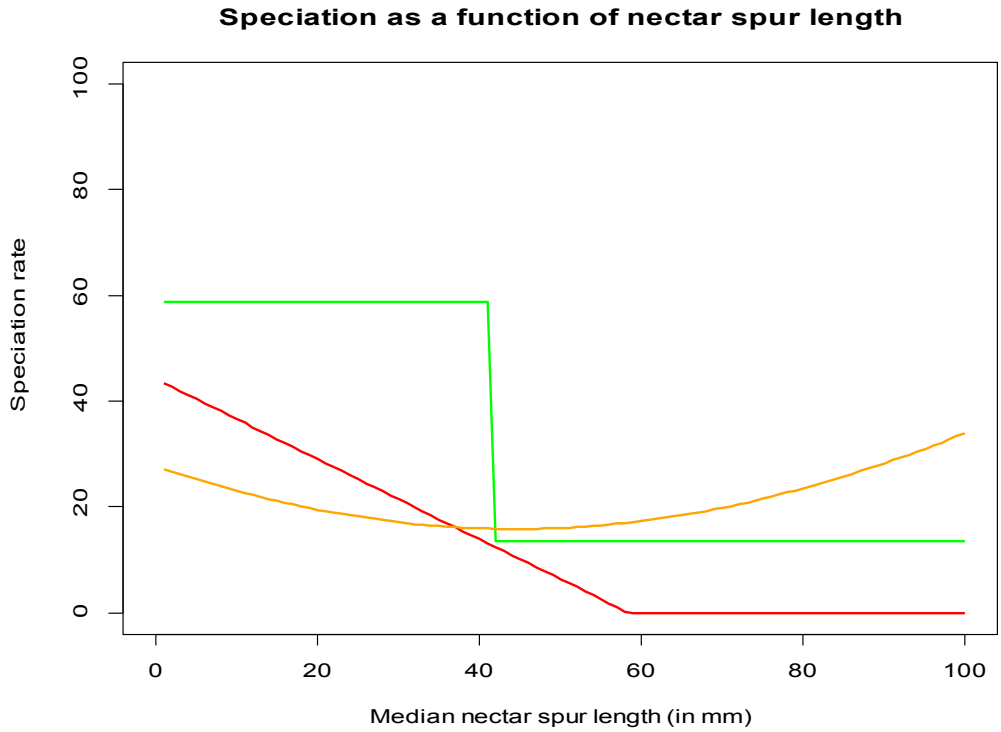

Figure 2.37. Speciation rate as a function of median nectar spur length for the A(BC)-topology.

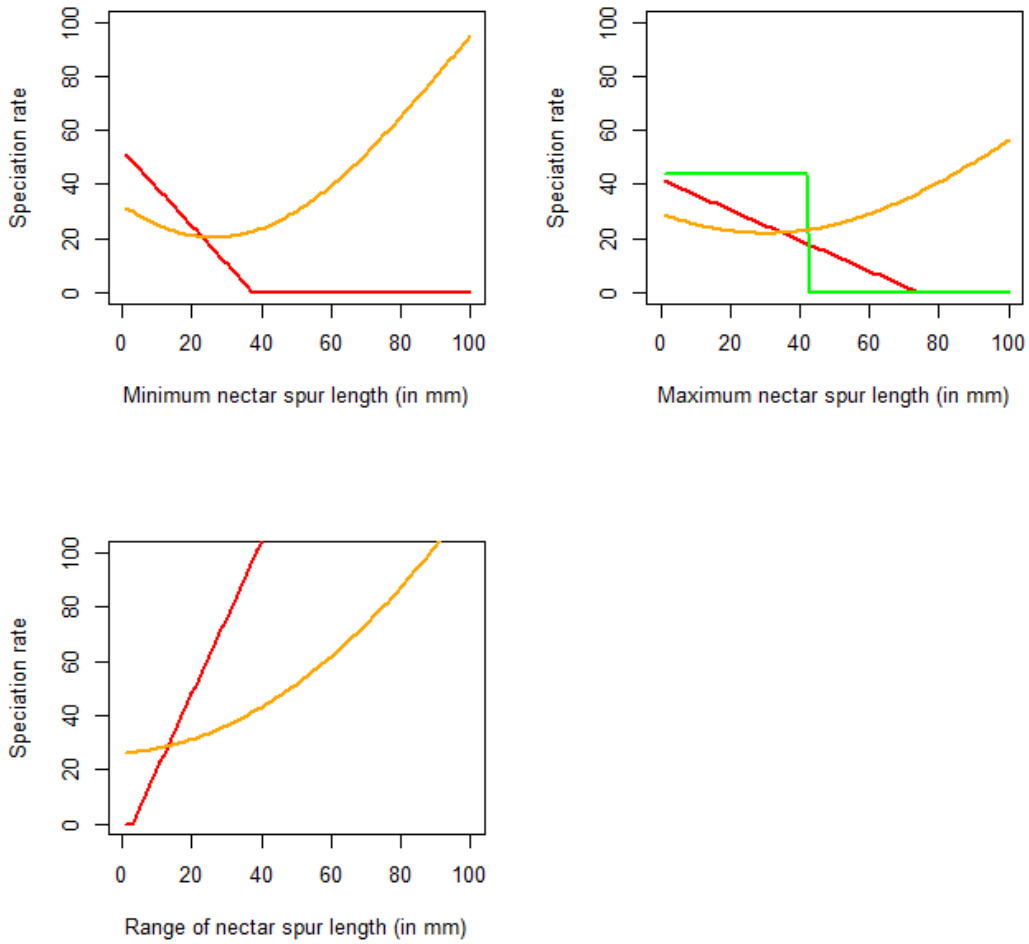

Figure 2.38. Speciation as a function of minimum, maximum and the range of nectar spur length in the A(BC)-phylogeny.

The pattern that emerges from using minimum and maximum spur lengths is quite similar to the pattern that was obtained used median spur length. Longer spurs tend to incorporate lower speciation rates according to the linear and sigmoid functions, whereas long spurs bring about higher speciation rates according to the parabolic function.

When using the range of spur lengths, things look different: a larger range brings about higher speciation rates, according to both the linear and the parabolic function. The parabolic model was chosen once as the best model, the linear model four times. The sigmoid model was never significant enough. Of all available models, 0.53 were chosen as a highly significant improvement over the null model of a constant speciation rate, not influenced by nectar spur length. The drift parameter for the parabolic functions was positive, whereas it was negative for the linear functions. This is the only case where the drift parameter was negative, indicating that there may not be a trend towards a larger range of nectar spur length.

(AB)C-topology

QuaSSE was run with median spur lengths on 15 trees of the (AB)C-topology. For every tree, the parabolic function was chosen as the best model. However, the other models were also a highly significant improvement over the null model; a fraction of 0.96 of all models had a probability value lower than 0.001. The results can be seen in figure 2.39, and are quite similar to the results from the conflicting topology. The biggest difference is that speciation rates at low nectar spur lengths tend to be a bit higher. However, speciation diminishes a lot quicker than for the A(BC)-topology; the linear function of speciation reaches 0 at a nectar spur length of 4 centimetres,

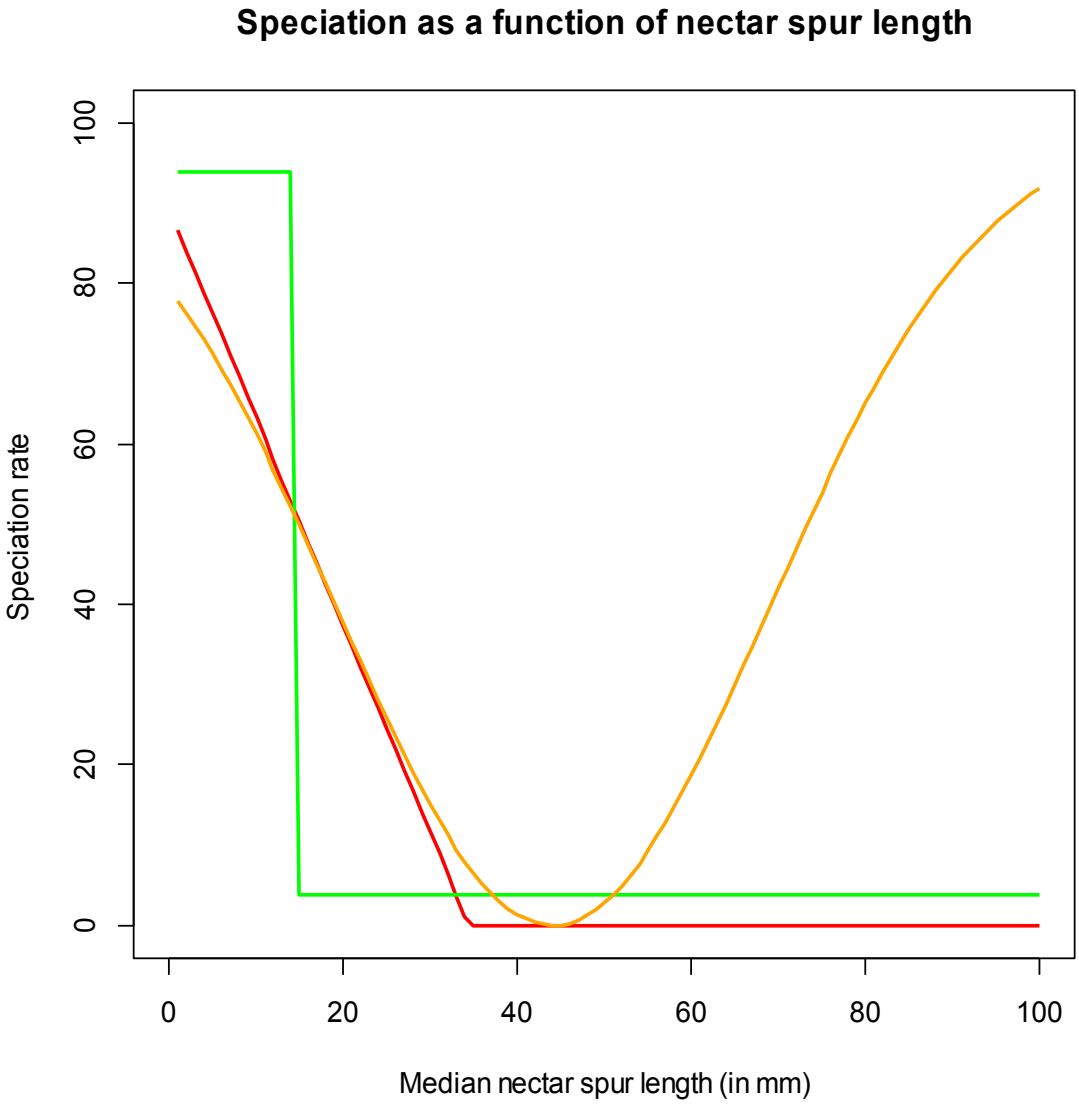

Figure 2.39. Speciation rate as a function of median nectar spur length for the A(BC)-topology.

whereas in the other topology, speciation did not become 0 until a nectar spur length of 6 centimetres. Similar effects can be seen for the other two functions.

A final function of QuaSSE is to fit speciation models that implement a diversification rate shift in the tree. Using MEDUSA (see the previous section), such a diversification rate shift was found at the base of the A-clades in the (AB)C-topology. Speciation models implementing this shift were fitted to 15 different trees, resulting in 2 speciation models per tree; one for the B- and C-clades,

and one for the A-clades. For each part of the tree, 2 different speciation models were used: a constant model, where speciation is not influenced by nectar spur length, and a linear model. This resulted in a total of four different combinations of these models. To save time, for both these models the drift parameter was set to 0.

The results were clear: for each tree, a model implementing a diversification rate shift provided a much better fit than a model using just one speciation function for the whole tree. However, it turned out to be impossible to distinguish between the different combinations of the two speciation functions; the models were virtually indiscernible based on their likelihood value and AIC scores. Therefore, all four combinations are equally preferred over the null model, and all four combinations are depicted in figure 2.40.

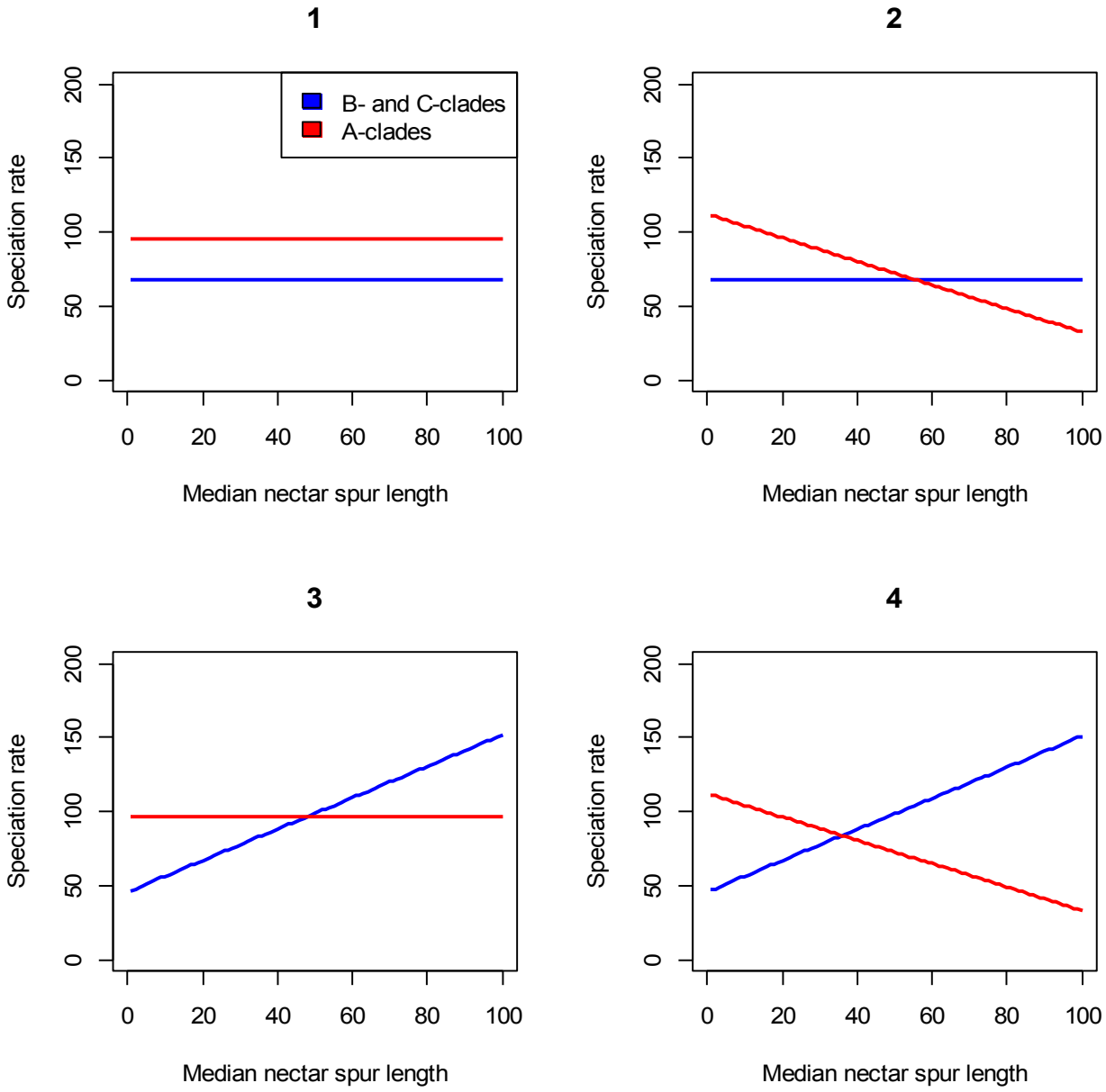

Figure 2.40. Speciation rate as a function of median nectar spur length for the A(BC)-topology, implementing a diversification rate shift at the base of the A-clades. The red line depicts speciation in the A-clades, the blue lines depicts speciation in the B- and C-clades.

Two important conclusions can be drawn based on figure 2.40. First of all, as can be seen in graph number 1, speciation rates in the A-clade are significantly higher than in the B- and C-clades. Considering the results from the MEDUSA-analysis, this does not come as a surprise. The second conclusion, however, is remarkable. As can be seen in graph number 2 and 4, longer nectar spur bring about lower speciation rates, but only in the A-clades (red line). In the B- and C-clades (blue line), longer nectar spurs in fact cause the speciation rate to increase (see graph number 3 and 4). This indicates that the B- and C-clades react in a completely different ways to an increase in nectar spur lengths than the A-clades.

Spur length evolution rate shifts

To identify shifts in nectar spur evolution, auteur was used to analyse 5 trees of both topologies. Two independent chains were run for 1000000 generations each, which led to effective sample sizes well over 100 for all parameters. The results were remarkably similar to the results from MEDUSA. For all 5 trees of the A(BC)-topology, small (containing 5 species at the most), random shifts were found all over the tree. No clear patterns could be distinguished regarding the direction or location of these shifts, just like in the MEDUSA analysis. However, the results for the (AB)C-topology were quite different; clear evidence for the probability of a shift in nectar spur evolution was found in 3 of the 5 trees. In all of these trees, the shift was found at the base of the A2-clades, and indicated that higher levels of spur evolution were found in this clade than in the rest of the genus. See figure 2.41 for the results of one of these trees.

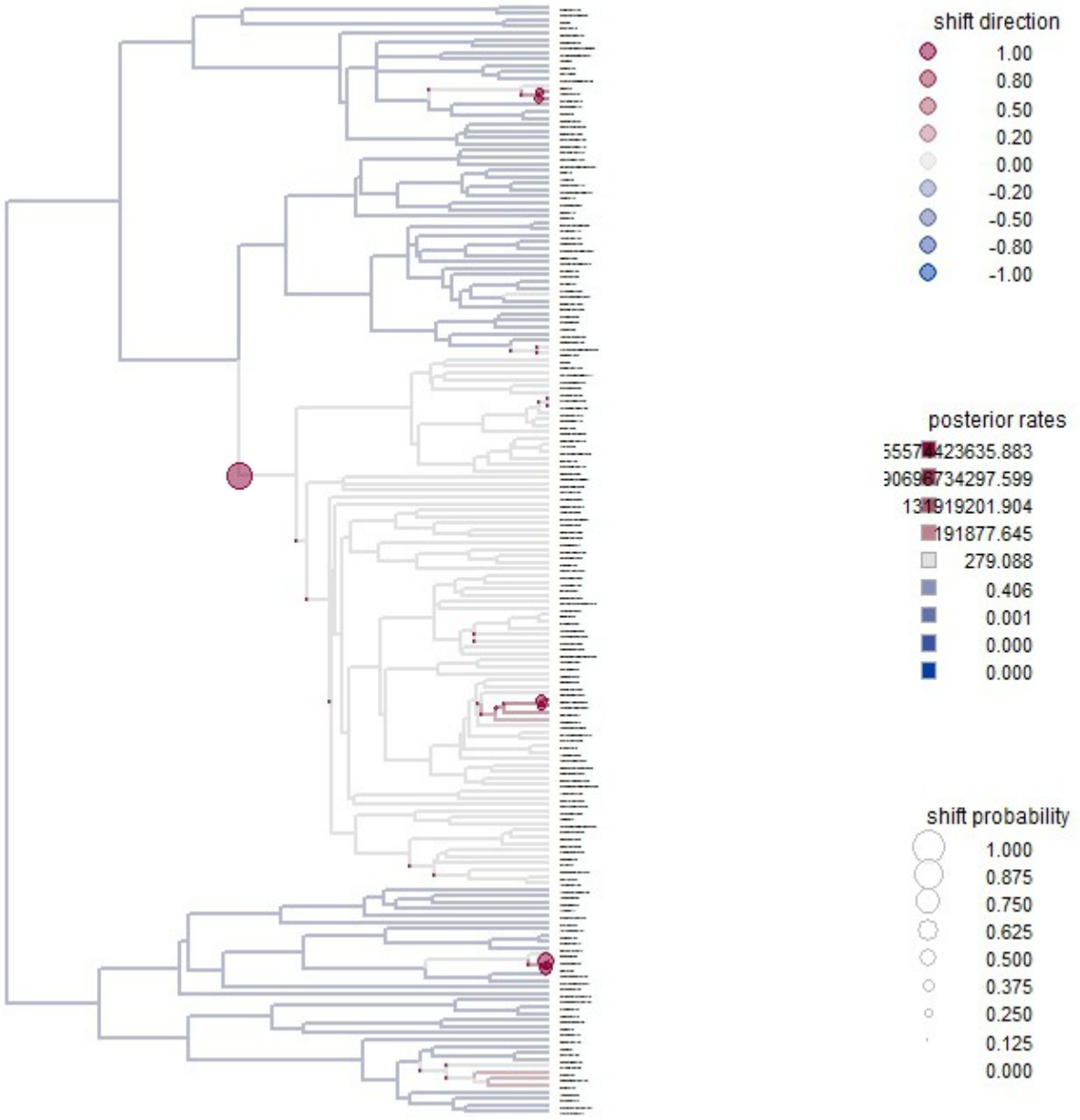

Figure 2.41. Locations and probabilities of nectar spur evolution rate shifts in the (AB)C-topology, showing evidence for higher levels of spur evolution in the A2-clades. Rate shifts are indicated with circles, where the diameter of the circle indicates the probability of the occurrence of the shift. The direction of the shift is marked by the colour of the circle, where red circles indicate shifts to higher rates of character evolution, and blue circles indicate lower rates. In the same way, the colour of the branches marks the rate of evolution.

Discussion

Trends towards longer spurs

The first hypothesis that will be discussed here covers trends in directionality: 'In the evolution of nectar spurs in *Pelargonium*, there was a trend towards longer spurs'. Of all research questions of this thesis, this one was easiest to answer: there is a clear trend to longer nectar spurs, as was pointed out by the QuaSSE analysis. This trend exists for minimum, median and maximum spur lengths, in both topologies. As FitzJohn (2010) writes in the original description of QuaSSE: "When the directional tendency opposed species selection, there was some power to detect the trend, but this power was never high for the parameter values explored". Therefore, the fact that QuaSSE does find such a trend (which opposes species selection, as will be discussed later) is a testimony to the strength of this trend; clearly it must be quite considerable, or else it would not have been picked up by QuaSSE. Furthermore, the SIMMAP analysis also indicated a trend towards longer nectar spurs, as (slightly) more transitions to larger spur lengths than to smaller spur lengths were calculated.

It is rather unfortunate that no other methods are available to test trends in the direction of evolution. As was discussed before, Pagel's Transformations (Pagel 1999 and 2002) are not capable of detecting such trends in ultrametric trees, and neither is the trend model of the "fitContinuous"-function (Geiger-package in R). Furthermore, when looking at nectar spur reconstructions in R for the nodes leading up to the most basal clades, no trend towards longer spurs can be distinguished. Therefore, the conclusion that there has been such a trend is solely based on the QuaSSE and SIMMAP analyses.

The existence of this trend leads to some interesting questions. For instance, how is it possible that there has been an evolutionary trend towards longer nectar spurs, while the nectar spurs of the majority of the species are relatively short (see figure 1.5, and Struck 1997)? For comparison: Whittal & Hodges (2007) found a trend towards longer nectar spurs in *Aquilegia*. They tested 30 monophyletic populations and species, of which only 4 had short nectar spurs. In *Pelargonium*, approximately 60% of all the species have nectar spurs which are shorter than 2 centimetres, and whole clades exist without any species with longer nectar spurs. One would expect more species with longer nectar spurs based on the presence of a trend towards these longer spurs.

A possible explanation of this phenomenon could be that there are drawbacks to evolving long spurs. These could be found in terms as larger energy costs to maintain longer nectar spurs, or higher risk of herbivory. Other possible trade-offs are that longer spurs tend to exclude more generalist pollinators (Hodges 1997, Borrell 2012) and increase the risk of nectar-robbing (Navarro & Medel 2009). In fact, the presence of one or more of such trade-offs to longer nectar spurs is deemed likely by the results of the QuaSSE analysis, which show that longer nectar spurs tend to incorporate lower speciation rates. This could encompass a whole range of processes, but one explanation is that individuals with longer nectar spurs have a lower probability to survive or reproduce. Of course, the exact cause of this drop in speciation rate does not become clear from the QuaSSE analysis, and experiments will need to be performed to find out the true reason why speciation rates are lower for species with longer nectar spurs. The QuaSSE analysis provides a description of the situation, not an explanation.

Spurs and speciation

The second hypothesis states: 'Pollinator-switches and clade-proliferation are linked in clades A2HLM and A2POC'. This hypothesis is difficult to test for several reasons. First of all, in this

thesis (and especially in this chapter) only nectar spur length data are available. No pollinator observations have been made, and the relationship between nectar spurs and pollinators remains unclear (see also chapter 3). Therefore, I will only be able to make statements regarding nectar spur evolution, rather than pollinator switches.
Secondly, the distinction between clade A2HLM and A2POC is rather trivial: both clades contain sections where much vegetative (for instance, transitions to a xerophytic lifestyle) and floral (nectar spur) evolution has occurred. Another complicating factor is that neither of these clades has been found to be completely monophyletic in the phylogenies reconstructed in this thesis. However, the complete A-clade, also containing clade A1, is a monophyletic clade, just as the two other major clades (clade B and C).
Therefore, this hypothesis will be divided into two subhypotheses: 'A correlation exists between nectar spur evolution and clade-proliferation' and 'Clade A has evolved in a different way than clades B and C'.

To start with the second subhypothesis: there are numerous indications that evolution in clade A has occurred in a different way than in clades B an C. In fact, most of the analyses performed in this chapter indicate a difference between clade A and the rest of the genus:

- the MEDUSA analysis hints at the presence of a diversification rate shift located at the base of the A-clades. However, this shift was not exactly clear, as it was only found in trees from the (AB)C-topology. The strongest evidence for this shift was found when analysing skeleton trees, where each clade is represented by a single species and the remaining species richness is described in a table. For these type of trees, a shift at the base of the A-clades was found in 86% of all analysed trees. However, when incorporating a complete tree, with all species, in only 23% of the trees a shift was found, and in trees of the A(BC)-topology, there was no evidence for the presence of such a diversification rate shift at all.
- similar to the results from MEDUSA, the auteur analysis points in the direction of higher nectar spur evolution in the A-clades. Once again, the evidence supporting this shift in spur evolution is rather weak. Shifts were only found in some trees of the A(BC)-topology.
- stronger evidence for a difference between clade A and clades B and C stems from the QuaSSE analysis. Models incorporating different effects of nectar spurs on speciation for the two different sections of the tree (A and B/C) provide a highly significant ($p < 0.001$) better fit than models that do not incorporate such a distinction between different sections of the tree. These models also show that the speciation rate is higher in the A-clades than in the rest of the tree. This is not the only support for the hypothesis that there is an evolutionary difference between clades A and B/C. According to QuaSSE, longer nectar spurs are correlated with higher rates of speciation in the B- and C-clades, but with lower speciation rates in the A-clade. This is clear evidence that in both sections of the tree, important differences exist regarding speciation and nectar spur evolution.
- additional support for this statement comes from the "fitContinuous"-function. Using this method, seven different models describing the evolution of a continuous character were fitted to median nectar spurs and trees from both topologies, after which the best model was selected based on its likelihood score. Clades A and C were usually best described by a model incorporating Pagel's λ, whereas a whole range of different models provided the best fit to clade B.
- finally, not only did different models fit to different parts of the trees, the parameters describing these models varied widely between the sections of the tree as well. λ-values, indicating the level of phylogenetic signal, tended to be higher for the A- and B-clades than for the C-clades, whereas κ-values, indicating whether character evolution occurs gradual or punctuational, were higher for the A-clade than for the rest of the tree.

All these results point in the direction that there are differences between clade A and clades B and C, both regarding the evolution of nectar spurs as regarding the rates of speciation (and extinction).

These results may not come as a complete surprise, considering that habitat, lifestyle and pollination strategies differ considerably between (some sections of) clade A and the rest of the genus *Pelargonium*. First of all, clade A is known as the Winter-rainfall clade (Bakker *et al.* 2005), as its species occur exclusively in the winter-rainfall region of the South African Cape (Linder 2003). Additionally, a large part of the species in this clade are adapted to the harsh, arid conditions of the succulent karoo, causing clades A2POC and A2HLM to be known as the xerophytic clade (Bakker *et al.* 2005). One strategy to deal with such xerophytic conditions is to adopt a geophytic lifestyle. All species of section *Hoarea* are geophytes, causing Bakker and colleagues to term it the geophytic clade. As most geophytic species are limited with respect to the amount of flowers they can produce, they have to adopt different pollination strategies, which is reflected by the impressive amount of floral evolution in *Hoarea* (Marais 1999). Other sections of the A-clade have also adopted different pollination strategies: the night-scented species of section *Polyactium* seem to focus purely on the attention of hawkmoths (Struck 1997), whereas the flowers of section *Otidia* are characterised by the presence of auricles at the base of the posterior petals, thereby blocking the opening of the nectar spur and limiting pollinators who seek to enter this spur (Becker & Albers, unpublished).
One could therefore conclude that there are several reasons and indications that evolution and speciation follow different patterns in clade A than in the rest of the phylogeny of *Pelargonium.*

The big question is: are these differences correlated with nectar spur evolution?
Judging on the MEDUSA and auteur analyses, one could conclude there is in fact such a correlation. As was seen before, the MEDUSA analysis shows support for higher speciation rates in clade A. Simultaneously, auteur shows an indication that there are higher levels of nectar spur evolution in clade A. Similar correlations between nectar spur evolution and clade diversification caused Hodges & Arnold (1995) to deem nectar spurs an evolutionary key innovation (Simpson 1953, Liem 1973). However, these rate shifts were only found in some trees of one topology. Therefore it is difficult to conclude there is a correlation between nectar spur evolution and the differences between clade A and clades B and C.

This is further corroborated by the results from QuaSSE. These results clearly show that longer nectar spurs are correlated with lower speciation rates in clade A. This is the opposite of what one would expect from an increase in spur length (Rabosky & McCune 2009), and also indicates that the increases in clade size of some sections in clade A were not (completely) caused by increased levels of nectar spur evolution. Interestingly, in other parts of the genus, longer spurs are correlated with higher levels of speciation. A possible explanation for this difference in reaction to nectar spur evolution may be found in the different lifestyles existing in the *Pelargonium* genus. Additional information is required before this difference can be fully understood and explained.

To conclude, one could argue there is no reason to reject the second subhypothesis, which states that 'Clade A has evolved in a different way than clades B and C'. This difference is indicated by various analyses. However, based only on the nectar spur measurements, no statements can be made regarding the causes of this difference.
The other subhypothesis, stating 'A correlation exists between nectar spur evolution and clade-proliferation', can be accepted nor rejected. It is clear that there is some kind of correlation between spur evolution and speciation, but the mechanisms behind this correlation, as well as its direction, remain disputed.

Nectar spurs and pollination

Apart from the main hypotheses that were tested, some observations are in need of further

explanation. As discussed before, one of the premises of this thesis is the idea that the nectar spur length of a *Pelargonium* species is an indication of its pollinator, which is stated in virtually every scientific study regarding *Pelargonium*, or nectar spurs in general. Furthermore, a general idea in nectar spur biology is that each species (or population) is pollinated by one, or at the most a few, pollinators. However, as was seen in the previous chapter, nectar spur distributions of individuals in a population are widely variable, and therefore these assumptions may be disputed. What does the evolution of nectar spurs suggest about the relation between spur and pollinator?

First of all, Pagel's λ shows that nectar spurs are characterised by an intermediate to relatively strong phylogenetic signal. This indicates that nectar spur lengths may be somewhat conserved throughout the genus, but can still vary considerably between two closely related species. Additionally, Pagel's κ shows that nectar spurs evolve in a punctuational way, which could be an indication of adaptive evolution (Pagel 2002). It also implies that spur length in *Pelargonium* evolution occurs according to the pollinator shift model (Whittall & Hodges 2007) rather than the evolutionary arms race model (Ennos 2008). As much nectar spur evolution occurs in each speciation event, it is possible that nectar spurs are involved in the speciation process.
Both analyses indicate there are selective constraints on the evolution of nectar spurs; it is clear they do not evolve in a gradual, unrestricted way. One explanation for this phenomenon is that nectar spurs evolve due to pollination pressure, but more information is needed before this can be fully tested.

Another indication that nectar spurs may be (up to a certain degree) a proxy for pollination syndrome are the QuaSSE analyses. They clearly show a correlation between nectar spur evolution and speciation, and the easiest explanation for this correlation is the existence of a close relationship between spurs and pollinators. This is partly confirmed by the analysis performed with the range of nectar spurs. As this range increased, speciation rates increased tremendously (see figure 2.38). If the premise regarding pollinators and nectar spurs is correct, there is an easy explanation for this: more variation in nectar spur length will attract many different pollinators, causing reproductive isolation between individuals or populations, and therefore speciation (Schluter 2001, Rabosky & McCune 2009). Interestingly, the range of spur length was the only measure of spur length without a clear trend towards longer spurs (or in this case, a larger range). Furthermore, Pagel's Transformations show that nectar spur range varies greatly throughout the genus; Pagel's λ is almost indistinguishable from 0, which means that the nectar spur range of a species does not depend on its place in the phylogeny.

In the previous chapter the hypothesis was proposed that several pollinators may exist per *Pelargonium* species and population. Judging on the relatively conserved minimum nectar spurs throughout the genus, it may be possible that individuals within a population adopt different pollination strategies; plants with small nectar spurs could attract generalist pollinators, such as bees, whereas plants with long nectar spurs attract more specialist pollinators, such as long-proboscid flies (Struck 1997). Pagel's Transformations corroborate this hypothesis. The high λ-values for minimum nectar spur length indicate that this measure is relatively conserved for all *Pelargonium* species; median and maximum spur length are much more variable between different species. Furthermore, the κ-values of minimum spur length approach 1, indicating gradual evolution, while the κ-values for median and maximum spur length are much lower, which is evidence for rapid changes and punctuational evolution. This pattern could be explained by different pollination strategies within species; short-spurred individuals are visited by generalist pollinators, which remain the same for many species throughout the genus, whereas longer-spurred individuals attract the attention of different pollinators.

Of course, if this is true, the big question is: why would different individuals within the same species focus on different pollinators? And if the pollinator fidelity hypothesis proves to be true, indicating that longer-spurred plants have a higher chance that their pollen will be delivered to individuals from the same species, why would not all individuals and species grow longer nectar spurs? A possible explanation could be that floral specialization may in fact have severe trade-offs; in *Ruellia* (Acanthaceae), Tripp & Manos (2008) showed that some specialized pollination systems may be an 'evolutionary dead-end'. In contrast, Davies *et al.* (2011) found that plant species traits of 10 Cape genera did not accurately predict extinction risk. To fully understand this process, much more information will be needed, such as pollinator observations, pollinator distributions, *Pelargonium* extinction records, and fossil data.

One of the original aims of this project was to reconstruction the nectar spur length of the hypothetical ancestral *Pelargonium*, as was done by Bakker *et al*. (2005). In that study, the ancestral spur length was reconstructed at 1-5 millimetres, which probably corresponds with pollination by bees. The ancestral reconstructions in this project paint a more complicated picture. The results from the "ace" analysis show that the most likely ancestral spur length was approximately 18 millimetres, which would correspond with pollination by a more specialist pollinator than bees, such as long-proboscid hovering flies. However, the results from the SIMMAP analysis were considerably more ambiguous, estimating high probabilities for wide ranges of ancestral nectar spur length, making it very difficult to draw any clear conclusions regarding ancestral pollination in *Pelargonium*.

Phylogenetic uncertainty and future recommendations

A final note regarding one of the premises of this thesis: 'Phylogenetic trees accurately portray the evolutionary history of the genus'. This premise was immediately challenged by the analyses of MrBayes, as two different topologies were created: the A(BC)-topology and the (AB)C-topology. Although there are indications that the (AB)C-topology probably predicts the 'true tree' (see Weng *et al.* 2012 as well as all previous reconstructions of the *Pelargonium* phylogeny), there was no reason to distinguish between the two versions of the *Pelargonium* topology, so all phylogenetic analyses were performed on trees from both topologies. This creates the interesting opportunity to compare the effect that the topology of a tree has on the results of an analysis, when all other variables (e.g., methods, models, characters) remain the same.

In general, two different patterns emerge:

- analyses such as ancestral state reconstructions, model fitting and inference of trait-dependent speciation rates do not (greatly) depend on the tree topology (when averaged over a range of trees from the same topology). Results were rather similar for both versions of the phylogenetic tree;
- however, for methods identifying shifts in evolution, be it species (MEDUSA) or character (auteur) evolution, the results varied greatly: in general, shifts were found in (some trees) of the (AB)C-topology, whereas no, or completely different, shifts were found in the A(BC)-topology. This difference is rather intriguing, and shows that plenty of thought and care should be given to the selection of the 'right' topology.

The results from the analyses performed in this thesis suggest three recommendations for future phylogenetic analyses:

1) Perform each analysis on as many trees as possible. For more or less every analysis performed, results, such as model parameters and reconstruction ancestral states, varied greatly per

tree, even if those trees were part of the same topology. This suggests that results from an analysis performed on one tree (even if that tree is a consensus tree) will be a severely limited interpretation of the true evolutionary processes that caused the observed patterns.

2) Similar to the recommendation from chapter 1, perform analyses using several measures of the character of interest (i.e., nectar spur length). Results may vary greatly, and give interesting insight in the processes of character evolution.

3) Finally, the results from this thesis suggest that the standard Brownian-motion model hardly ever predicts character evolution in an accurate way. Other models of evolution should be considered and tested before any further analyses are performed.

Chapter 3:
Nectar spurs and pollinators

Introduction

This thesis treats nectar spur lengths of *Pelargonium* species as a 'one-dimensional proxy' for their pollinator syndrome; by analysing the distribution and evolution of nectar spur length, the aim is to test hypotheses regarding pollinator-shifts and diversification. The previous chapters did not question this approach, but in this chapter I will address this premise.

To analyse whether nectar spur lengths accurately predict pollinators, pollinator and nectar spur information is combined in one graph. As an example of other floral characters that could influence pollinator-choice, flower colour is also optimised over a range of trees, to see if any trends in colour evolution and pollinator-shifts can be identified.

Another question that will be addressed is whether speciation in *Pelargonium* can be said to be pollinator-driven, or if other processes could also explain the putative correlation between diversification and nectar spur length evolution found in chapter 2. I will discuss whether this question can be fully answered with the available data.

As the above questions will outline some limitations in this thesis, several recommendations for future research on pollination in *Pelargonium* are made.

Materials and Methods

Relation pollinators and nectar spur lengths

To analyse the relation between nectar spur lengths and pollinators, information regarding the two was obtained from different sources. All *Pelargonium* species with known pollinators, as described by Struck in 1997, were incorporated in the pre-existing nectar spur length dataset. For all these 70 species, nectar spur lengths were already obtained previously, either from field or herbarium observations, or from recordings in the scientific literature. For easy analysis, the results were plotted with different colours per pollinator.

Ancestral flower colours

Similar to the ancestral nectar spur reconstructions performed in chapter 2, for all nodes in the *Pelargonium* phylogeny ancestral flower colours were reconstructed. As this was mainly done for graphical purposes, rather than real hypothesis testing, the analysis was not performed in R, but in Mesquite (Maddison & Maddison 2011). A set of trees of the (AB)C-topology, consisting of 1 50% consensus tree and 200 randomly-chosen trees, was loaded into Mesquite, together with a character set depicting the flower colours of the *Pelargonium* species. For this purpose, flower colour was divided into 7 discrete classes (white-pink, yellow, maroon, yellow-brown-green, pink-red, red, and green), based on descriptions in the primary scientific literature. Ancestral states were analysed using the "Trace Characer Over Trees" function. The reconstruction method was Unordered Parsimony (indicating that each change is equally costly, and therefore equally likely). Reconstructed nodes were depicted on the consensus tree, to take both phylogenetic and reconstruction uncertainty into account.

Results

Pollinators and nectar spur lengths

To analyse the relation between nectar spur length and pollinator, information was obtained from several sources, and plotted with different colours for easy analysis. The results are in figure 3.1.

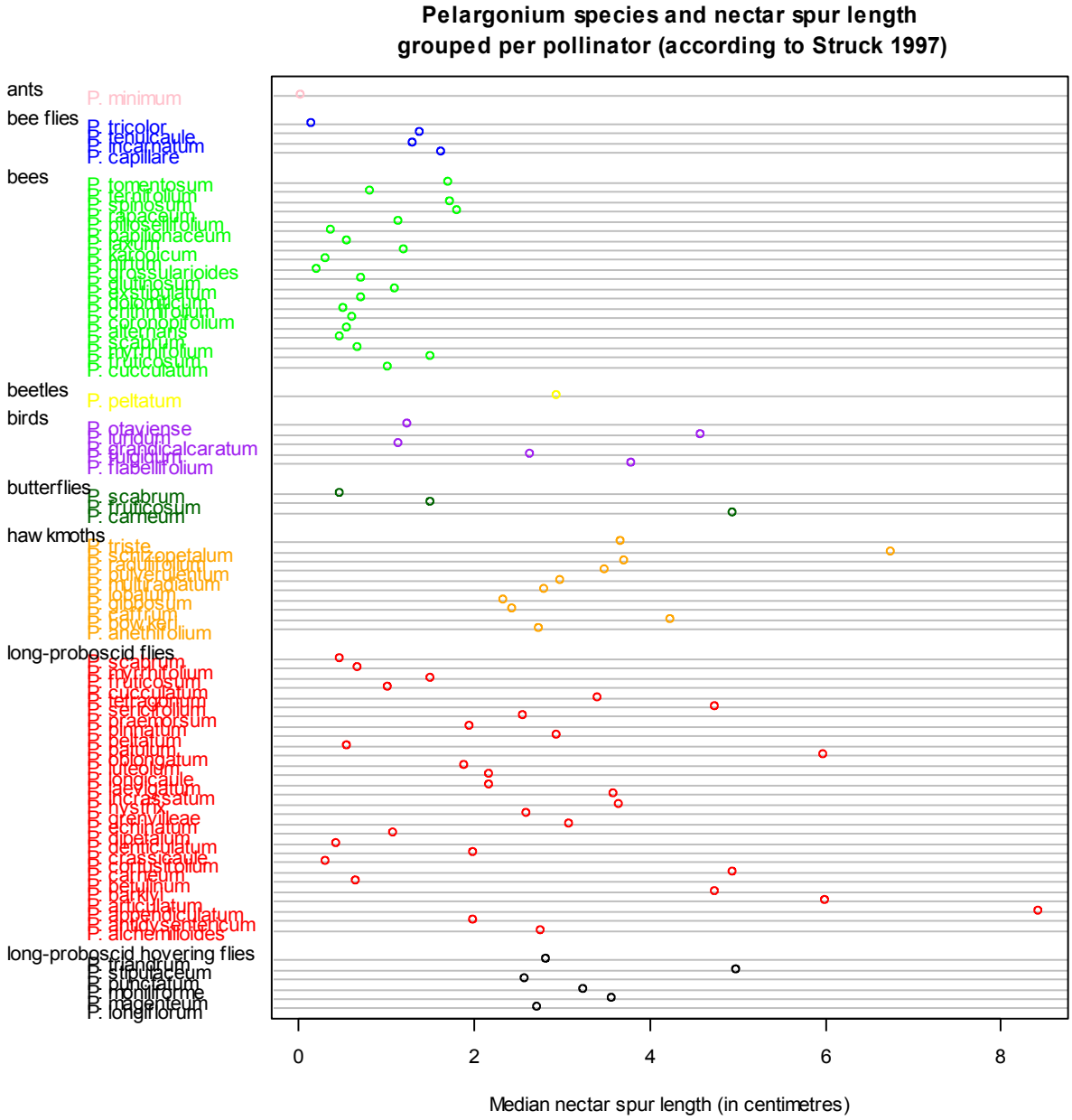

Figure 3.1. Median nectar spur lengths for *Pelargonium* species grouped per pollinator, according to Struck (1997).

Struck (1997) identified 7 different pollinators: ants, bee flies, bees, beetles, birds, butterflies, hawk moths, long-proboscid flies and long-proboscid hovering flies. As can be seen, *Pelargonium* species with different pollinators tend to be characterised by different nectar spur lengths. Species with pollinators such as ants, bees, and bee flies have rather short nectar spurs, rarely exceeding 2

centimetres. In contrast, species pollinated by hawk moths or long-proboscid flies have nectar spurs up to 8 centimetres long. However, for most pollinator groups, and especially for species with longer nectar spurs, much variation in spur length exists between species pollinated by the same animal. For example, some species pollinated by long-proboscid flies have nectar spurs which are, based purely on their length, indistinguishable from species pollinated by bees. Furthermore, the nectar spurs lengths of almost every pollinator group have considerable overlap with nectar spurs from other groups. It is clear that, at least for some species, a relation exists between pollinators and nectar spur lengths, but identifying a pollinator purely based on the nectar spur length of a species is extremely difficult, if not impossible.
It is important to note that this analysis only takes into account median nectar spur lengths. As was seen in chapter 1, some *Pelargonium* species have high levels of spur length variation; as this is not taken into account in figure 3.1, this figure provides a simplified overview of a much more complex situation.

Ancestral flower colours

The reconstructed ancestral flower colours, plotted on the 50% consensus tree of the (AB)C-topology, can be seen in figure 3.2.

Based on this simple analysis, it is difficult to discover trends regarding flower colour evolution in *Pelargonium.* It is clear that some sections have specific colours (such as the yellow-brown coloured flowers of section *Polyactium*), but the vast majority of the species is simply coloured the characteristic pink-white. Selecting and appointing pollination syndromes based on flower colour (and nectar spur length) is difficult; for a proper analysis of pollination syndromes, one would also need, among other characteristics, information on nectar guides, nectar composition, scent, and the length and the position of the stamens. Unfortunately this information is not readily available, making appointing pollination syndromes rather difficult, if not impossible, to do.

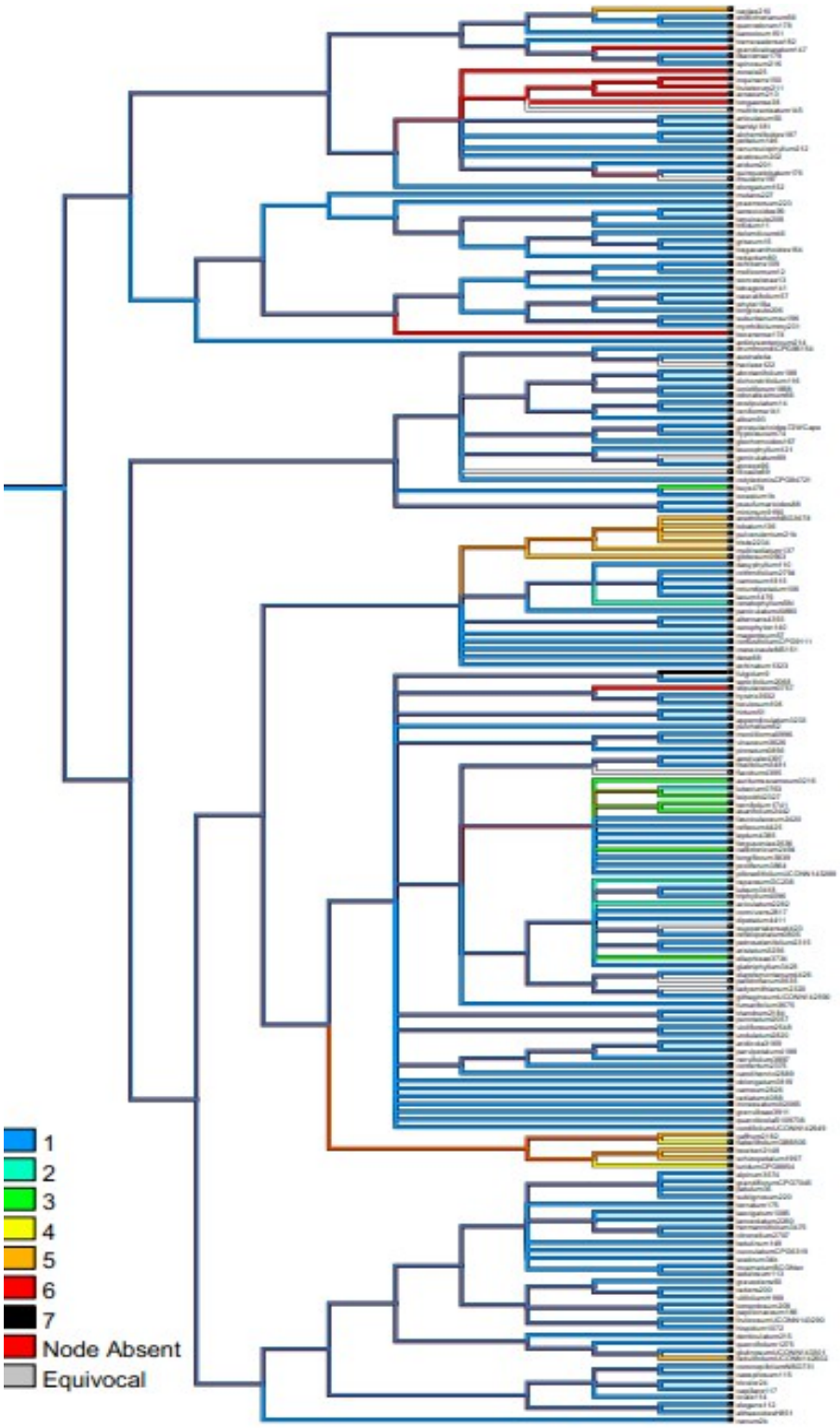

Figure 3.2. Ancestral flower colours plotted on the 50% consensus tree of the (AB)C-topology. Legend: 1 = white-pink, 2 = yellow, 3 = maroon, 4 = yellow-brown-green, 5 = pink-red, 6 = red, 7 = green.

Discussion

The relation between nectar spurs and pollinators

The most important premise of this thesis, 'Nectar spur length is a 'one-dimensional proxy' for pollinator syndrome', is probably also the most easily-disputed statement made in this thesis. Based on the analyses performed in this chapter, it seems clear that nectar spur length and pollinator are related: groups of *Pelargonium* species pollinated by different pollinators tend to be distinguished (in broad ways) by their nectar spur lengths. However, extensive variation in nectar spur length exists within such pollinator groups, and there are large levels of overlap between the nectar spur lengths of different groups. Identifying a pollinator purely based on the nectar spur length of a species is simply not possible (Struck 1997).

Adding flower colour to the dataset does not provide much clarification. Some patterns can be distinguished, but as the vast majority of the species is known for its cream-white coloured flowers, no clear patterns can be distinguished, let alone pollination syndromes. To properly analyse pollination syndromes, information about many more floral characters is needed: nectar composition (Struck 1997), length and position of the stamens (Marais 1999), scent of the flowers, structure of the plant (Hanley *et al*. 2009), and nectar guides (Hansen *et al.* 2012), to name but a few. If all this information would be available, the first steps could be made to start a proper identification of pollination syndromes in *Pelargonium*.

Even with information about all those plant characteristics available, care should be taken when identifying pollination syndromes. As Ollerton and colleagues (2009) showed in a world-wide analysis of putative pollination syndromes, floral characters and their pollinators, the most-common pollinator could not be predicted for two-thirds of all their studies species. The authors therefore suggest caution when appointing floral characters and pollination syndromes. Furthermore, it has been shown that, at least in the irid *Tritoniopsis revoluta*, pollinators other than the 'main pollinator' predicted by the pollination syndrome actually provided a more important function in the pollination of plants of this species (de Merxem *et al.* 2009). As this is a South African species with long corrolla tubes, there is no reason to assume this process does not occur in some *Pelargonium* species.

There is a simple explanation for this: the pollinator of a certain plant species depends on a whole range of processes and characteristics other than simply its pollination syndrome (Van der Niet & Johnson 2012). Interactions with non-pollinating animals such as florivores (Strauss & Whittall 2006), for example, may influence the floral characteristics of a plant, and therefore its relation to pollinators. Similarly, the evolution of chemical (Armbruster 1997) or physical (Hanley *et al.* 2009) defence mechanisms may open up possibilities for new pollinators. It also seems likely to assume that nectar robbers may have influence on the evolution of floral characters and plant-pollinator interactions (Navarro & Medel 2009). Of course, the availability and density of different pollinators will also have a large impact on which animal will act as the main pollinator (Johnson 2010). This clearly suggests that extreme care should be take when identifying pollination syndromes. Doing so purely based on morphological characters, without analysing any actual pollinator observations, will almost always result in the erroneous identification of the primary pollinators (Van der Niet & Johnson 2012).

It may be clear the pollinator - nectar spur premise does not withstand close scrutiny. The results from this thesis are therefore mainly applicable to nectar spur evolution and distribution, rather than pollinator shifts.

Pollinator-driven speciation

One of the main questions in evolutionary biology is why new species are formed; what is the exact driver of the initial reproductive isolation and, therefore, speciation (Schluter 2001)? For plants in the South African Cape, considerable debate exists about the exact methods of speciation. Some studies (e.g., Van der Niet & Johnson 2008, Johnson 2010, Johnson & Anderson 2010) suggest that speciation is pollinator-driven, whereas other studies (e.g., Van der Niet *et al.* 2006, Schnitzler *et al.* 2011) conclude that shifts in soil-type use are the main drivers of speciation. Can this thesis shed any light on this discussion?

The results from the QuaSSE analysis suggest that the evolution of nectar spur length has a significant correlation with speciation rates in *Pelargonium*. Furthermore, the co-occurrence of diversification and spur length evolution rate shifts seem to indicate that nectar spur evolution (and therefore presumably pollinators) play an important role in speciation in *Pelargonium*. However, it is important to realise that this kind of conclusions can not be drawn based purely on the available information (Losos 2011).

The reason for that is simply that correlation does not automatically mean causation. The (partial) correlation between diversification and nectar spur length evolution found in this thesis could be explained by the occurrence of pollinator-driven speciation in *Pelargonium* (Whittall & Hodges 2007). However, at least two other explanations are also possible (Van der Niet & Johnson 2012). First of all, nectar spur evolution (and thereby floral divergence) may be the method of reinforcement of already existing reproductive isolation, rather than the first driver of such isolation (Schluter 2001). This could happen when two recently-diverged subpopulations come into secondary contact and hybrids of these two subpopulations have a lower fitness than pure-breeds, in which case reproductive isolation is actively selected for. This process is called reinforcement, and could also describe the found correlation between spur evolution and diversification. A second alternative explanation is that nectar spur divergence between species is caused by character displacement (Armbruster & Machhala 2009): competition for pollinators between species (which are not necessarily related) may drive divergent evolution of floral characters, such as nectar spurs, to attract different pollinators and thereby avoid competition (Rodríguez-Gironés & Llandres 2008, Rodríguez-Gironés & Santamaría 2010). The difference between these three processes is their timing: if speciation is driven by pollination, pollinator-switches will occur before speciation events. If floral divergence evolves to reinforce already existing reproductive isolation, floral evolution will take place during or shortly after speciation. Finally, pollinator shifts because of character displacement will occur after speciation. With no information regarding the occurrence and exact timing of pollinator shifts, and the complete lack of information about other ecological characters, such as soil-type, it is impossible to conclude whether or not speciation in *Pelargonium* is pollinator-driven (Armbruster & Machhala 2009, Losos 2011, Van der Niet & Johnson 2012). All that can be said is that the rather punctuational evolution of nectar spur length seems to suggest an important role for nectar spurs in fitness, and possibly speciation, but the exact relation between nectar spur evolution, fitness and speciation can not be determined based on the available information.

Recommendations for future research

This chapter identifies some crucial shortcomings of this thesis. As both time and money are limited in this project, I will not be able to overcome these shortcomings. However, if there is an opportunity to perform more research on nectar spur length evolution and pollinators in *Pelargonium*, the following actions are recommended:

1) Perform pollinator observations. The most crucial shortcoming of this thesis is the complete lack of pollinator observations. The most recent inquiry into pollination in *Pelargonium* stems from 15 years ago (Struck 1997) and is far from complete. Before any questions regarding the role of pollinators in *Pelargonium* can be answered, extensive field observations will have to be made. These observations will preferably need to be performed on the population level, for as many as populations and species as possible. This will allow answering questions that need to be addressed before any further research can take place, such as questions regarding the exact relation between species and pollinators, the number of pollinators per species, and the selectional pressure on both plant and pollinator.

2) Identify pollination syndromes. When sufficient pollinator observations are available, a first step can be made to identifying pollination syndromes. The best approach would be to perform a complete morphological analysis of as many species as possible (e.g., Van der Niet *et al.* 2011). Characters of interest include flower colour, length and position of stamens, nectar spur length and width, effective nectar spur length (rather than absolute spur length, as measured in this thesis), scent, structure of plant, size of flower, composition of nectar, and many others. When pollination syndromes are accurately known for the majority of the *Pelargonium* species, pollination syndromes for the remaining species can be predicted based on their floral characters. Using pollinator observations it can be tested whether these predictions are accurate or not. Once all the pollination syndromes are known, they can be mapped onto a phylogeny, so questions regarding directionality and evolutionary dead-ends can be answered (Whittal & Hodges 2007).

3) Identify which genes are responsible for nectar spur evolution. It would be highly interesting to know whether only a few nectar spur-influencing genes are available, or if the length of a nectar spur is determined by a whole range of genes. Similarly, an assessment of heritability and the measure of genetic versus environmental variation could be made (Hartl 2000).

4) Collect other ecological information about *Pelargonium*. To assess whether speciation in *Pelargonium* is driven by pollination or by something else, more information is needed. A complete database, containing information about soil type use, climatic preferences, fire resistance, etcetera, is needed before this question can be addressed.

Once this has all been done, hypotheses regarding speciation and pollination in *Pelargonium* can be truly tested.

Conclusions

- Nectar spur length distribution in *Pelargonium* species is characterised by large levels of variation within and between certain species, populations, and individuals. However, for some other species, the levels of spur length variation are much lower. These varying levels of spur length variation may be explained by the Geographic Mosaic Theory of Coevolution (Pauw *et al.* 2008).

- There is a clear evolutionary trend towards longer nectar spurs, as was indicated by the results from QuaSSE (and possibly SIMMAP). Unfortunately no other methods are available to test for evolutionary trends in ultrametric trees.

- Nectar spur evolution and speciation rates are correlated, which is an indication of species selection (Rabosky & McCune 2009); however, the direction and mechanism of this correlation are unclear. In the A-clade of *Pelargonium*, longer nectar spurs are correlated with lower speciation rates, whereas in the B and C-clades the opposite pattern was found.

- Nectar spur length evolution in *Pelargonium* is characterised by a moderate to strong phylogenetic signal and occurs in a punctuational way, which is an indication that spur length evolution occurs according to the pollinator shift model (Whittal & Hodges 2007).

- No clear conclusions can be drawn regarding the nectar spur length of the hypothetical ancestral *Pelargonium*. Results from the "ace"-function indicate an ancestral spur length of approximately 18 millimetres, whereas results from SIMMAP consider an ancestral spur length between 0 and 25 millimetres as the most likely. Based on these reconstructions it is not possible to infer an ancestral pollinator.

- Minimum nectar spur lengths seem more conserved over species and populations than maximum spur lengths. This is corroborated three observations; 1) minimum nectar spurs have a higher phylogenetic signal than median and maximum spur lengths, 2) minimum nectar spurs evolve in a more gradual way than median and maximum spur lengths, and 3) minimum nectar spur length is poorly correlated with the range of nectar spur length of a species. These findings may indicate that individuals with shorter spurs are visited by the same generalist pollinators, but individuals with longer spurs may attract the attention of more specialist pollinators.

- Clear differences exist between the A-clade of *Pelargonium* and the B and C-clades, regarding evolution (higher speciation rates, different models of spur length evolution), ecology (different life forms and survival strategies), and pollination (nectar spur length evolution has a different effect on speciation).

- Two different *Pelargonium* topologies were inferred in this project; the A(BC)-topology and the (AB)C-topology. Based on other studies, the second topology seems the most likely.

- The results from this project show that using different phylogenetic trees, methods, and measures of spur length for answering the same questions may lead to fundamentally different results. Therefore I strongly recommend to always take into account phylogenetic, character and method/model uncertainty.

- Nectar spur length is not a perfect proxy for pollination syndrome. Before any hypotheses about speciation and pollination in *Pelargonium* can be truly tested, extensive pollinator observations are needed.

Acknowledgements

People

I am very grateful to Freek Bakker and Timo van der Niet for providing excellent, extensive and patient supervision of my thesis project. I am greatly indebted to Stuart Hall for his willingness to share his astonishing knowledge of the South African flora and his dedication to find and collect as many *Pelargonium* species as possible. Bettie Marais and Leanne Dreyer provided invaluable information about *Pelargonium*, pollination, and fieldwork in South Africa, and made Stellenbosch University feel like home to me. Terry Trinder-Smith is kindly thanked for providing extensive support while working in the Bolus Herbarium, and Terry Hedderson kindly allowed me to give a presentation at Department of Botany of the University of Cape Town. Furthermore I would like to thank Robin van Velzen for spending so much of his time to instruct me in using R. Lars Chatrou, Martijn Staats, Nienke Groendijk-Wilders, Lotte Joosten, Ria Vrielink-van Ginkel, Theodoor Heijerman, Ronald van den Berg, André van Proosdij, Wilma Twigt, Jos van der Maesen, and Jan Wieringa are thanked for providing me with useful comments, information and materials for my project. Joop Ringelberg's advices on using Linux and Mac computers and on programming in general have been extremely useful. Finally I would like to thank everyone at the Biosystematics Group, and especially all the people from The Island (Mathijs, Erik-Jan, Rutger, Jeike, Mark, Rens, Pulcherie, Ximena, Sjoerd, and Erik) for making this group such a pleasant place to have been working on my thesis.

Organisations

South African National Parks and CapeNature, as well as all the respective national parks and conservation areas, have kindly allowed me to perform field work on their lands, for which I am very obliged to them. The staff of Elandsberg Private Nature Reserve, which is usually closed for public, allowed me to work on their property, for which I am very grateful. I would also like to thank the respective staffs of the Bolus Herbarium, the Compton Herbarium, the Stellenbosch University Herbarium, and the Herbarium Vadense for allowing me to use their collections, just like the Department of Botany of the University of Cape Town for allowing me to give a presentation to their group. Furthermore I would like to thank the staff of the Botanischer Garten of the University of Münster for giving me access to their collection of *Pelargonium* plants.

Finally I am very grateful to the Alberta Mennega Stichting, Stichting Moabi, Taxon Ondersteuning, and the Biosystematics Group for providing me with financial support to complete the fieldwork of this thesis.

References

Aikaike, H (1974) **A new look at the statistical model identification**. IEEE Transactions on Automatic Control, Volume 19, Issue 6, pages 716–723.

Albers, F, Gibby, M & Austmann, M (1992). **A reappraisal of *Pelargonium* sect. *Ligularia* (Geraniaceae)**. Plant Systematic Evolution, Volume 179, pages 257-276.

Alfaro, ME, Santini, F, Brock, C, Alamillo, H, Dornburg, A, Rabosky, DL, Carnevale, G & Harmon, LJ (2009) **Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates**. Proceedings of the National Academy of Sciences of the United States of America, Volume 106, Issue 32.

Anderson, B, Terblanche JS & Ellis, AG (2010) **Predictable patterns of trait mismatches between interacting plants and insects**. BMC Evolutionary Biology, 10:204.

Armbruster, WS (1997) **Exaptations link evolution of plant-herbivore and plant-pollinator interactions: a phylogenetic inquiry**. Ecology 78, 1661–1672.

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

Bakker, FT, Culham, A, Hettiarachi, P, Touloumenidou, T & Gibby, M (2004) **Phylogeny of *Pelargonium* (Geraniaceae) based on DNA sequences from three genomes**. Taxon, Volume 53, pages 17-28.

Bakker, FT, Culham, A, Marais, EM & Gibby, M (2005) **Nested radiation in Cape *Pelargonium***. Pages. 75–100 *in* Bakker, FT, Chartrou, LW, Gravendeel, B, & Pielser, PB eds. Plant species-level systematics: new perspectives on pattern and process. A. R. G. Ganter Verlag K. G., Ruggell, Liechtstein.

Bakker, FT, Breman F & Merckx, V (2006) **DNA sequence evolution in fast evolving mitochondrial DNA *nad*1 exons in Geraniaceae and Plantaginaceae**. Taxon, Volume 55, Issue 4, pages 887-896.

Bannikov, AF & Carnevale, G (2011) **Enigmatic spiny-rayed fish from the Eocene of Monte Bolca, Italy**. Geological Journal 46, 1, pages 52-62.

Bartolucci, F, Scaccia, L, & A. Mira, A (2006) **Efficient Bayes factor estimates from the reversible jump output**. Biometrika 93: 41–52.

Becker, M & Albers, F (Unpublished) **Radiation and pollinator shift in *Pelargonium alternans* J.C.Wendl. (Geraniaceae)**.

Barraclough, TG (2006) **What can phylogenetics tell us about speciation in the Cape flora?** Diversity and Distributions, Volume 12, pages 21–26.

Bloch, D, Erhardt, A (2008) **Selection toward shorter flowers by butterflies whose probosces are shorter than floral tubes**. Ecology, Volume 89, pages 2453–60.

Bollback, JP (2006) **SIMMAP: Stochastic character mapping of discrete traits on phylogenies**. BMC Bioinformatics, Volume 7, Issue 88.

Bollback, JP (2009) **SIMMAP Documentation - Morphology/Standard Priors**. http://www.simmap.com/pgs/priors.html. Last visited on 26-10-2012.

Bond, P & Goldblatt, P (1984) **Plants of the Cape Flora**. Journal of South African Botany, Suppl. 13: 1±455.

Born, J, Linder, HP & Desmet, PG (2006) **The Greater Cape Floristic Region**. Journal of Biogeography, Volume 33, pages 1–16.

Borrell, BJ (2005) **Long Tongues and Loose Niches: Evolution of Euglossine Bees and Their Nectar Flowers**. Biotropica 37(4): 664–669.

Bowman, AW & Azzalini, A (2010) **R package 'sm': nonparametric smoothing methods** (version 2.2-4) URL http://www.stats.gla.ac.uk/~adrian/sm, http://azzalini.stat.unipd.it/Book_sm.

Box, MS, Dodsworth, S, Rudall, PJ, Bateman, RM & Glover, BJ (2011) **Characterization of Linaria KNOX genes suggests a role in petal-spur development**. Plant Journal 68 (4), pages 703-714.

Brown JW, FitzJohn RG, Alfaro ME & Harmon, LJ (2012) **turboMEDUSA: Modelling Evolutionary Diversification Using Stepwise AIC**. R package version 0.93-2.9.1.

Burbrink, FT, Ruane, S & Pyron, RA (2012) **When are adaptive radiations replicated in areas? Ecological opportunity and unexceptional diversification in West Indian dipsadine snakes (Colubridae: Alsophiini)**. Journal of Biogeography, 39 (3), pages 465-475.

Burnham KP & Anderson DR (2003) **Model Selection and Multimodel Inference, A Practical Information-Theoretic Approach** (Springer, New York).

Butler, MA & King, AA (2004) **Phylogenetic comparative analysis: a modeling approach for adaptive evolution**. American Naturalist, Volume 164, pages 683-695.

Cantalapiedra, JL, Fernández, MH & Morales, J (2011) **Biomic Specialization and Speciation Rates in Ruminants (Cetartiodactyla, Mammalia): A Test of the Resource-Use Hypothesis at the Global Scale**. PLoS ONE, 6 (12), art. no. e28749.

Capellini, I, Barton, RA, McNamara, P, Preston, BT & Nunn, CL (2008) **Phylogenetic analysis of the ecology and evolution of mammalian sleep**. Evolution, Volume 62, Issue 7, pages 1764-1776.

Cardini, A & Elton, S (2008) **Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons**. Biological Journal of the Linnean Society, Vol. 93, Issue 4, pages 813-834.

Carlson, JE, Holsinger, KE & Prunier, R (2011) P**lant responses to climate in the Cape Floristic Region of South Africa: evidence for adaptive differentiation in the Proteaceae**. Evolution, Volume 65, Issue 1, pages 108-124.

Casewell, NR, Huttley, GA & Wüster, W (2012) **Dynamic evolution of venom proteins in squamate reptiles**. Nature Communications, Volume 3, Article number 1066.

Claramunt, S, Derryberry, EP, Remsen Jr, JV & Brumfield, RT (2011). **High dispersal ability inhibits speciation in a continental radiation of passerine birds**. Proceedings of the Royal Society B: Biological Sciences 297, pages 1567-1574.

Combs, JK & Pauw, A (2009) **Preliminary evidence that the long-proboscid fly, *Philoliche gulosa*, pollinates *Disa karooica* and its proposed Batesian model *Pelargonium stipulaceum***. South African Journal of Botany 75, pages 757– 761.

Covain, R, Dray, S, Fisch-Muller, S & Montoya-Burgos, JI (2007) **Assessing phylogenetic dependence of morphological traits using co-inertia prior to investigate characterevolution in Loricariinae catfishes**. Molecular Phylogenetics and Evoltuion, Volume 46, Issue 3, pages 986-1002.

Cunningham, CW, Omland, KE & Oakley, TH (1998) **Reconstructing ancestral character states: A critical reappraisal**. Trends in Ecology and Evolution, 13: 361–366.

Cunningham, CW (1999) **Some limitations of ancestral character-state reconstruction when testing evolutionary hypotheses**. Systematic Biology, 48: 665–674.

Darwin, CR (1859) **On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life**. London, UK: John Murray.

Darwin, CR (1862). **On the various contrivances by which British and foreign orchids are fertilized by insects**, facsimile edn. London, UK: JohnMurray.

Darwin C (1876) **The effects of cross—and self-fertilization in the vegetable kingdom**. London, UK: John Murray.

Darwin, CR (1877) **The various contrivances by which orchids are fertilised by insects**. London, UK: John Murray.

Davies, TJ, Smith, GF, Bellstedt, DU, Boatwright, JS, Bytebier, B, Cowling, RM, Forest, F, Harmon, LJ, Muasya, AM, Schrire, BD, Steenkamp, Y, Van der Bank, M & Savolainen, V (2011) **Extinction Risk and Diversification Are Linked in a Plant Biodiversity Hotspot**. PLoS Biology, Volume 9, Issue 5.

Delpino, F (1873–1874) **Ulteriori osservazioni e considerazioni sulla dicogamia nel regno vegetale**. Atti della Societa` Italiana di Scienze Naturali 16: 151–349, 17: 266–407.

deMenocal, PB (2004) **African climate change and faunal evolution during the Pliocene-**Pleistocene. Earth and Planetary Science Letters, Volume 220, pages 3-24.

de Merxem, DG, Borremans, B, de Jäger, ML, Johnson, T, Jooste, M, Ros, P, Zenni, RD, Ellis, AG & Anderson, B (2009) **The importance of flower visitors not predicted by floral syndromes**. South African Journal of Botany 75, 660– 667.

De Wet, LAR, Barker, NP & Peter, CI (2008) **The long and the short of gene flow and**

reproductive isolation: Inter- Simple Sequence Repeat (ISSR) markers support the recognition of two floral forms in *Pelargonium reniforme* (Geraniaceae). Biochemical Systematics and Ecology, Volume 36, pages 684–690.

Dobzhansky, T (1951) **Genetics and the Origin of Species** (3rd edn), Columbia University Press.

Eastman, JM, Alfaro, ME, Joyce, P, Hipp, AL & Harmon, LJ (2011) **A novel comparative method for identifying shifts in the rate of character evolution on trees**. Evolution (early view), published online 24 August 2011.

Eastman, JM, Alfaro, ME, Hipp, AL, Joyce, P, Pennell, M & Harmon, LJ (2012). **auteur: Bayesian sampler of the trait-evolutionary process**. R package version 0.12.0118. http://CRAN.R-project.org/package=auteur

Ekman, S, Andersen, HL & Wedin, M (2008) **The Limitations of Ancestral State Reconstruction and the Evolution of the Ascus in the Lecanorales (Lichenized Ascomycota)**. Systematic Biology, 57(1): 141–156.

Eldenas, PK & Linder, HP (2000) **Congruence and complementarity of morphological and trnL-trnF sequence data and the phylogeny of the African Restionaceae**. Systematic Botany, Volume 25, pages 692-707.

Ennos, RA (2008) **Floral evolution – Spurred on by pollinators**. Heredity , 100, 3–4.

Estes, S & Arnold, SJ (2007) **Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales**. Am. Nat. 169: 227–244.

Fægri, K & Van der Pijl, L (1979) **The principles of pollination ecology**. Pergamon Press, Oxford

Felsenstein, J (1985) **Phylogenies and the comparative method**. Am. Nat.125, 1–15.

Felsenstein, J (1988) **Phylogenies and quantitative characters**. Annu. Rev. Ecol. Syst.19, 4 45–471.

FitzJohn, RG, Maddison, WP & Otto, SP (2009) **Estimating Trait-Dependent Speciation and Extinction Rates from Incompletely Resolved Phylogenies**. Syst. Biol. 58(6): 595-611.

FitzJohn, RG (2010) **Quantitative Traits and Diversification**. Syst. Biol. 59(6): 619-633.

FitzJohn, RG, with GeoSSE, ClaSSE by Goldberg, EE, and BiSSE-ness by Magnuson-Ford, K (2012) **Diversitree: comparative phylogenetic analyses of diversification**. R package version 0.9-3. http://CRAN.R-project.org/package=diversitree.

Frumhoff, PC & Reeve, HK (1994) **Using phylogenies to test hypotheses of adaptation : a critique of some current proposals**. Evolution 48, 172–180.

Gaschen, B, Taylor, J, Yusim, K, Foley, B, Gao, F, Lang, D, Novitsky, V, Haynes, B, Hahn, BH, Bhattacharya, T & Korber, B (2002) **Diversity considerations in HIV-1 vaccine selection**. Science 296, 2354–2360.

Gibby, M, Hinnah, S, Marais, Em & Albers, F (1996) **Cytological variation and evolution within *Pelargonium* section *Hoarea*** (Geraniaceae). Plant Systematic Evolution, Volume 203, pages 111-142.

Gilbert, GS & Webb, CO (2007) **Phylogenetic signal in plant pathogen-host range**. Proceedings of the National Academy of the United States of America, Vol. 104, no. 12, pages 4979-4983.

Givnish, TJ & Sytsma, A (1997) **Molecular evolution and adaptive radiation**. Cambridge, UK: Cambridge University Press.

Goldblatt, P (1978) **An analysis of the flora of Southern Africa: its characteristics, relationships, and origins**. Annals of the Missouri Botanical Garden, Volume 65, pages 369–436.

Goldblatt, P & Manning, JC (1999) **The long-proboscid fly pollination system in Gladiolus (Iridaceae)**. Annals of the Missouri Botanical Garden, Vol. 86, No. 3, pages 758:774.

Goldblatt, P & Manning, JC (2000) **The long-proboscid fly pollination system in southern Africa**. Annals of the Missouri Botanical Gardens 87, pages 146-170.

Goldblatt, P & Manning, J (2000). **Cape Plants. A Conspectus of the Cape Flora of South Africa**. National Botanical Institute, Pretoria.

Goldblatt, P & Manning, JC (2006) **Radiation of pollination systems in the Iridaceae of sub-Saharan Africa**. Annals of Botany, Volume 97, pages 317–344.

Gould, SJ (1997) **Cope's rule as psychological artefact**. Nature, Volume 385, pages 199–200.

Grant, V (1949) **Pollination systems as isolating mechanisms in angiosperms**. Evolution, Volume 3, pages 82–97.

Grant, V & Grant, KA (1965) **Flower pollination in the Phloxfamily**. New York, NY: Columbia University Press.

Guisingera, MM, Kuehl, JV, Boore, JL & Jansena, RK (2008) **Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions**. PNAS, Volume 105, Number 47, pages 18424-18429.

Hanley, ME, Lamont, BB & Armbruster, SW (2009) **Pollination and plant defence traits co-vary in Western Australian Hakeas**. New Phytol. 182, 251–260.

Hansen, DM, Van der Niet, T & Johnson, SD (2012) **Floral signposts: testing the significance of visual 'nectar guides' for pollinator behaviour and plant fitness**. Proceedings of the Royal Society B 279 , 634–639.

Hardy, CR & Linder, HP (2005) **Intraspecific variability and timing in ancestral ecology reconstruction: a test case from the Cape flora**. Systematic Biology, 54, pages 299–316.

Hardy, CR (2006) **Reconstructing ancestral ecologies:challenges and possible solutions**. Diversity and Distributions, Volume 12, pages 7–19.

Harmon *et al.* **Early burst**. [Paper in revision.]

Harmon L, Weir J, Brock C, Glor R, Challenger W & Hunt, G (2009). **geiger: Analysis of evolutionary diversification**. R package version 1.3-1. http://CRAN.R-project.org/package=geiger.

Hartl, DL (2000) **A primer of population genetics**. Third edition. Sinauer Associates, Incorporated.

Hastings, WK (1970) **Monte-Carlo sampling methods using Markov chains and their applications**. Biometrika 57: 97–109.

Hedges, SB, Parker, PH, Sibley, CG & Kumar, S (1996) **Continental breakup and the ordinal diversification of birds and mammals**. Nature, Volume 381.

Hodges, Sa & Arnold, ML (1995) **Spurring plant diversification: are floral nectar spurs a key innovation?** Proceedings: Biological Sciences, 262, pages 343-348.

Hodges, SA (1997) **Floral nectar spurs and diversification**. International Journal of Plant Sciences, Vol. 158, No. 6, Supplement: Morphology and Evolution of Flowers (Nov., 1997), pages S81-S88.

Hodges, SA & Whittall, JB (2008) **Floral evolution – One-sided evolution or two? A reply to Ennos**. Heredity, Volume 100, pages 541-541.

Huelsenbeck, JP, Rannala, B & Masly, JP (2000) **Accommodating Phylogenetic Uncertainty in Evolutionary Studies**. Science, Vol. 288, No. 5475.

Huelsenbeck, JP & Ronquist, F (2001) **MRBAYES: Bayesian inference of phylogeny**. Bioinformatics 17:754–755.

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

Huelsenbeck, JP, Larget, B & Alfaro, ME (2004) **Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo**. Molecular Biology and Evolution 21, pages 1123:1133.

Janzen, DH (1980) **When is it coevolution?** Evolution, Volume 34, pages 611–612.

Johnson, MTJ, FitzJohn, RG, Smith, SD, Rausher, MD & Otto, SP (2011) **Loss of sexual recombination and segregation is associated with increased diversification in evening primroses**. Evolution 65-11: 3230–3240.

Johnson, SD & Steiner, KE (1997) **Long-tongued fly pollination and evolution of floral spur length in the Disa draconis complex (Orchidaceae)**. Evolution, Volume 51, Number 1, pages 45-53.

Johnson, SD, Linder, HP & Steiner, KE (1998) **Phylogeny and radiation of pollination systems in Disa (Orchidaceae)**. American Journal of Botany, Volume 85, pages 402–411.

Johnson, SD (2010) **The pollination niche and its role in the diversification and maintenance of**

the southern African flora. Philosophical Transactions of the Royal Society B, 2010 365, 499-516.

Johnson, SD & Anderson, B (2010) **Coevolution Between Food-Rewarding Flowers and Their Pollinators**. Evolution: Education and Outreach, Volume 3, Number 1, pages 32-39.

Johnson, SD, Peter, CI, Ellis, AG, Boberg, E, Botes, C & Van der Niet, T (2011) **Diverse pollination systems of the twin-spurred orchid genus *Satyrium* in African grasslands**. Plant Systematic Evolution, Volume 292, pages 95-103.

Jones, CS, Bakker, FT, Schlichting, CD & Nicotra, AB (2009) **Leaf shape evolution in the south african genus *Pelargonium* L' Hér. (Geraniaceae)**. Evolution, Volume 63, Issue 2, pages 479–497.

Kramer, EM & Hodges, SA (2010) **Aquilegia as a model system for the evolution and ecology of petals**. Philosophical Transactions of the Royal Society B: Biological Sciences 365 (1539), pages 477-490.

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, Volume 40, pages 637-656.

Kelley, NB, Near, TJ & Alonzo, H (2012) **Diversification of egg-deposition behaviour and the evolution of male parental care in darters (Teleostei: Percidae: Etheostomatinae)**. Journal of Evolutionary Biology, Volume 25, pages 836-846.

Liem, KF **(**1973**) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws.** Systematic Zoology, Volume 22, pages 425- 441.

Linder, HP (2003) **The radiation of the Cape flora, southern Africa**. Biological Reviews, Volume 78, pages 597-638.

Linder, HP & Hardy, CR (2004) **Evolution of the species-rich Cape flora**. Philosophical Transactions: Biological Sciences, Volume 359, Number 1450, Plant Phylogeny and the Origin of Major Biomes, pages 1623-1632.

Little, KJ, Dieringer, G & Romano, M (2005) **Pollination ecology, genetic diversity and selection on nectar spur length in Platanthera lacera (Orchidaceae)**. Plant Species Biology 20 (3), pages 183-190.

Loehrlein, MM & Craig, R (2001) **History and Culture of Regal Pelargonium**. HortTechnology, Volume 11, Issue 2, pages 289-296.

Losos, JB (2011) **Seeing the forest for the trees: the limitations of phylogenies in comparative biology**. American Naturalist 177, 709–727.

Maddison, WP, Midford, PE & Otto, SP (2007) **Estimating a Binary Character's Effect on Speciation and Extinction**. Syst. Biol. 56(5): 701-710.

Maddison, WP & Maddison, DR (2011) **Mesquite: A modular system for evolutionary analysis.** Version 2.75. http://mesquiteproject.org

Marais, EM (1981) **Die Taksonomiese Waarde van die Blom van Pelargonium** (Geraniaceae), Master thesis.

Marais, EM (1996) ***Pelargonium tripalmatum*** **(Geraniaceae), a new species from the richtersveld**. South African Journal of Botany, Volume 62, Issue 5, pages 225-227.

Marais, EM (1999) **Ideas on the pollination of *Pelargonium,* section *Hoarea* (Geraniaceae)**. In: Timberlake, J. & Kativu, S. (eds.). African Plants: Biodiversity, Taxonomy and uses. pp. 105–112. Royal Botanic Gardens, Kew.

Marloth, R (1908). **Das Kapland, insonderheit das Reich der Kapflora, das Waldgebiet und die Karroo, pflanzengeographisch dargestelt**. In Wissenschaffliche Ergebnisse der deutschen Tiefsee-Expedition auf dem Dampfer 'Valdivia' 1898–1899, Volume 2, Part 3 (ed. C. Chun), pp. 1–427. Gustav Fischer, Jena.

Mayr, E (1942) **Systematics and the Origin of Species**, Columbia University Press.

Metropolis, N, Rosenbluth, AW, Rosenbluth, MN, Teller, AH & Teller, E (1953) **Equation of state calculations by fast computing machines**. J. Chemical Phys. 21: 1087–1092.

Micheneau, C, Carlsward, BS, Fay, MF, Bytebier, B, Pailler, T & Chase,MW (2008) **Phylogenetics and biogeography of Mascarene angraecoidorchids (Vandeae, Orchidaceae)**. Molecular Phylogenetics and Evolution, Volume 46, pages 908–22.

Micheneau, C, Johnson, SD & Fay, MF (2009) **Orchid pollination: From Darwin to the present day**. Botanical Journal of the Linnean Society, Volume 161, pages 1–19.

Moore, DS & McCabe, GP (2009) **Introduction to the Practice of Statistics, fifth edition**. W.H. Freeman & Company, New York and Basingstoke.

Müller, H (1873) **Proboscis capable of sucking the nectar of *Angraecum sesquipedale***. Nature, 8, 223.

Müller, H (1882) **Weitere Beobachtungen über Befruchtung der Blumen durch Insekten, Teil III**. Verhandlungen des naturhistorischen Vereines der preussischen Rheinlande und Westfalens. Neununddreissigster Jahrgang. Vierte Folge: 9. Jahrgang, 1–104.

Nagy, LG, Házi, J, Szappanos, B, Koscubé, S, Bálint, B, Rákhely, G, Vágvölgyi, C & Papp, T (2012) **The evolution of defense mechanisms correlate with the explosive diversification of autodigesting coprinellus mushrooms (agaricales, fungi)**. Systematic Biology 61, 4, pages 595-607.

Navarro, L & Medel, R (2009) **Relationship between floral tube length and nectar robbing in *Duranta erecta* L. (Verbenaceae)**. Biological Journal of the Linnean Society, 96, 392–398.

Oberlander, KC, Dreyer, LL, & Esler, KJ (2002) **Biogeography of *Oxalis* (Oxalidaceae) in South Africa: a preliminary study**. Bothalia, Volume 3, pages 97-100.

Ollerton, J, Alarcón, R, Waser, NM, Price, MV, Watts, S, Cranmer,L, Hingston, A, Peter, CI & John Rotenberry (2009) **A global test of the pollination syndrome hypothesis**. Annals of Botany,

Volume 103, Issue 9, pages 1471-480.

O'Meara, BC, Ane, C, Sanderson, MJ & Wainwright, PC (2006) **Testing for different rates of continuous trait evolution using likelihood**. Evolution, Volume 60, pages 922–933.

Orr, MR & Smith, TB (1998) **Ecology and speciation**. Trends in Ecology and Evolution, Volume 13, pages 502–506.

Otto, SP & Whitton, J (2000) **Polyploid incidence and evolution**. Annual Review of Genetics, Volume 34, pages 401–437.

Pagel, M (1999) **Inferring evolutionary processes from phylogenies**. Zoologica Scripta, Volume 26, pages 331-348.

Pagel, M (2002) **Modelling the evolution of continuously varying characters on phylogenetic trees - The case of Hominid cranial capacity**. In Morphology, Shape and Phylogeny. Edited by Norman MacLeod and Peter L . Forey. CRC Press 2002. Pages 269–286.

Pagel, M, Meade, A & Barker, D (2004) **Bayesian Estimation of Ancestral Character States on Phylogenies**. Systematic Biology 53(5):673–684.

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

Pauw ,A, Stofberg ,J & Waterman, RJ (2008) **Flies and flowers in Darwin's race**. Evolution, Volume 63, Issue 1, pages 268-279.

Plummer, M, Best, N, Cowles, K & Vines, K (2006) **CODA: Convergence Diagnosis and Output Analysis for MCMC**, R News, vol 6, pages 7-11.

R Development Core Team (2012) **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rabosky, DL & McCune, AR (2009) **Reinventing species selection with molecular phylogenies**. Trends in Ecology and Evolution, Volume 25, number 2, pages 68-74.

Rabosky, DL (2010) **Extinction rates should not be estimated from molecular phylogenies**. Evolution 64-6: 1816–1824.

Rabosky, DL & Adams, DC (2012) **Rates of morphological evolution are correlated with species richness in salamanders**. Evolution 66, 6, pages 1807-1818.

Rainey, PB & Travisano, M (1998) **Adaptive radiation in a heterogeneous environment**. Nature, Volume 394.

Rambaut A & Drummond AJ (2007) **Tracer v1.5**, 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, Volume 3, Issue 2, pages 217-223.

Revel, LJ, Mahler, DL, Peres-Neto, PR & Redelings, BD (2012) **A new phylogenetic method for identifying exceptional phenotypic diversification**. Evolution, Volume 66, Issue 1, pages 135-146.

Rodríguez-Gironés, MA & Santamaría, L (2007) **Resource competition, character displacement, and the evolution of deep corolla** tubes. The American Naturalist, Volume 170, Number 3, pages 455-464.

Rodríguez-Gironés, MA & Llandres, AL (2008) **Resource Competition Triggers the Co-Evolution of Long Tongues and Deep Corolla Tubes**. PLoS ONE, Volume 3, Issue 8, e2992.

Rodríguez-Gironés, MA & Santamaría, L (2010) **How Foraging Behaviour and Resource Partitioning Can Drive the Evolution of Flowers and the Structure of Pollination Networks**. The Open Ecology Journal, 2010, Volume 3.

Rohlf, FJ (2010a) **tpsDig, digitize landmarks and outlines**. Version 2.16. Department of Ecology and Evolution, State University of New York at Stony Brook.

Rohlf, FJ (2010b) **tpsUtil, file utility program**. Version 1.47. Department of Ecology and Evolution, State University of New York at Stony Brook.

Ronquist, F (2004) **Bayesian inference of character evolution**. TRENDS in Ecology and Evolution, Vol.19, No.9.

Ronquist, F, Huelsenbeck, JP & Teslenko, M (2011) **Draft MrBayes version 3.2 Manual: Tutorials and Model Summaries**. Online publication, http://mrbayes.sourceforge.net/manual.php (last visited on October 2, 2012).

Rymer, PD, Manning, JC, Goldblatt, P, Powell, MP & Savolainen, V. (2010) **Evidence of recent and continuous speciation in a biodiversity hotspot: a population genetic approach in southern African gladioli (Gladiolus; Iridaceae)**. Molecular Ecology, Volume 19, Issue 21, pages 4765-4782.

Sargent, RD (2004) **Floral symmetry affects speciation rates in angiosperms**. Proceedings of the Royal Society London B, Issue 271, pages 603-608.

Schluter, D, Price, T, Mooers, AØ & Ludwig, D (1997) **Likelihood of ancestor states in adaptive radiation**. Evolution 51:1699–1711.

Schluter, D (1998) **Ecological causes of speciation**. In *Endless Forms: Species and Speciation* (Howard, D. and Berlocher, S., eds),pp. 114–129, Oxford University Press.

Schluter, D (2001) **Ecology and the origin of species**. TRENDS in Ecology & Evolution, Volume16, Number 7, pages 372-380.

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, Volume 60, Issue 3, pages 343-357.

Schwarz, MP, Bull, NJ & Cooper, SJB (2003) **Molecular phylogenetics of allodapine bees, with implications for the evolution of sociality and progressive rearing**. Systematic Biology, 52, 1–14.

Siepielski, AM, DiBattist, JD & Carlson, SM (2009) **It's about time: the temporal dynamics of phenotypic selection in the wild**. Ecology Letters, Volume 12, Issue 11, pages 1261-1276.

Simpson, GG (1944) **Tempo and Mode in Evolution**. Columbia University Press, New York.

Simpson, GG (1953) **The major features of evolution.** New York: Columbia University Press.

Stang, M (2007) **The structure of flower visitation webs – How morphology and abundance effect interaction patterns between flowers and flower visitors**. PhD Thesis, Leiden University.

Strauss, SY & Whittall, JB (2006) **Non-pollinator agents of selection on floral traits**. In The Ecology and Evolution of Flowers (Harder, LD & Barrett, SCH, eds), pp. 120–138, Oxford University Press.

Stebbins, GL (1970) **Adaptive radiation of reproductive characteristics in angiosperms. I. Pollination mechanisms**. Annual Review of Ecological Systems, Volume 1, pages 307–326.

Struck, M (1997) **Floral divergence and convergence in the genus *Pelargonium (Geraniaceae)* in southern Africa: ecological and evolutionary considerations**. Plant Systematic Evolution, Volume 208, pages 71-97.

Takhtajan, A (1986) **Floristic Regions of the World**. University of California Press, Berkeley.

Thibert-Plante, X & Hendry, AP **Factors influencing progress toward sympatric speciation**. Journal of Evolutionary Biology, Volume 24, Issue 10, pages 2186-2196.

Thornton, JW, Need, E & Crews, D (2003) **Resurrecting the ancestral steroid receptor: ancient origin of estrogen signalling**. Science, 301, 1714–1717

Thunberg, CP (1813) **Flora Capensis**. Edman, Uppsala.

Toulomenidou, T, Bakker, FT, Marais, EM & Albers, F (2004) **Chromosomal evolution interpreted from the rDNA ITS phylogeny for *Pelargonium* sect. *Hoarea (Geraniaceae).*** In: Thiede J, Veste M, Jurgens N, Metzing D ed(s). Sukkulentenforschung in Afrika. Succulent plant research in Africa. Festschrift Prof. Hans-Dieter Ihlenfeldt. (Schumannia 4; Biodiversity and ecology 2). Hamburg: Deutsche Kakteen-Gesellschaft pp.93-106.

Tripp, EA& Manos, PS (2008) **Is floral specialization an evolutionary dead-end? Pollination system transitions in *Ruellia* (Acanthaceae)**. Evolution 62-7: 1712–1737.

Van der Niet, T, Johnson, SD & Linder, HP (2006) **Macroevolutionary data suggest a role for reinforcement in pollination system shifts**. Evolution, Volume 60, Issue 8, pages 1596-1601.

Van der Niet, T & Johnson, SD (2008) **Patterns of plant speciation in the Cape floristic region**. Molecular Phylogenetics and Evolution, Volume 51, Issue 1, pages 85-93.

Van der Niet, T, Liltved, WR & Johnson, SD (2011) **More than meets the eye: A morphological and phylogenetic comparison of long-spurred, white-flowered satyrium species (Orchidaceae) in South Africa**. Botanical Journal of the Linnean Society, Volume 166, Issue 4, Pages 417-430.

Van der Niet, T & Johnson, ST (2012) **Phylogenetic evidence for pollinator-driven diversification of angiosperms**. Trends in Ecology and Evolution, Vol. 27, No. 6.

Van der Niet, T & Bakker, FT (Unpublished) **Mechanisms of pollinator-driven speciation in Pelargonium (Geraniaceae) inferred from combining ecological traits with a species tree**.

Van der Walt, JJA (1977) **Pelargoniums of Southern Africa, vol. 1.** Purnell, Cape Town, South Africa.

Van der Walt, JJA & Vorster, PJ (1981) **Pelargoniums of Southern Africa, vol. 2.** Juta, Cape Town, South Africa.

Van der Walt, JJA & Vorster, PJ (1988) **Pelargoniums of Southern Africa, vol. 3.** Jutta, Capetown.

Venditti, C & Pagel, M (2008) **Evolution in fits and starts**. Biologist, Volume 55 Number 3.

Verboom, GA, Linder, HP & Stock, WD (2004) **Testing the adaptive nature of radiation: growth form and life history divergence in the African grass genus *Ehrharta* (Poaceae: Ehrhartoideae)**. American Journal of Botany, Volume 91, pages 1364-1370.

Verboom, GA, Archibald, JK, Bakker, FT, Bellstedt, DU, Conrad, F, Dreyer, LL, Forest, F, Galley, C, Goldblatt, P, Henning, JF, Mummenhoff, K, Linder, HP, Muasya, AM, Oberlander, KC, Savolainen, V, Snijmanm, DA, Van der Niet, T & Nowell, TL (2009) **Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both?** Molecular Phylogenetics and Evolution, Volume 51, pages 44–53.

Via, S. (2001) **Sympatric speciation in animals: the ugly duckling grows up**. Trends in Ecology and Evolution, Volume 16, pages 381–390.

Warren, BH, Bakker FT, Bellstedt, DU, Bytebier, B, Claßen-Bockhoff, R, Dreyer, LL, Edwards, D, Forest, F, Galley, C, Hardy, CR, Linder, HP, Muasya, AM, Mummenhoff, K, Oberlander, KC, Quint, M, Richardson, JE, Savolainen, V, Schrire, BD, Van der Niet, T, Verboom, GA, Yesson, C & Hawkins, JA (2011) **Consistent phenological shifts in the making of a biodiversity hotspot: the Cape flora**. BMC Evolutionary Biology 2011, pages 11-39.

Weinreich, DM, Delaney, NF, DePristo, MA, Hartl, DL (2006) **Darwinian Evolution Can Follow Only Very Few Mutational Paths to Fitter Proteins**. Science, Volume 312, pages 111-114.

Weller, SG & Sakai, AK (1999) **Using phylogenetic approaches for analysis of plant breeding system evolution.**
Annual Review of Ecological Systems, Volume 30, pages 167–199.

Weng, ML, Ruhlman, TA, Gibby, M & Jansen, RK (2012) **Phylogeny, rate variation, and genome size evolution of Pelargonium (Geraniaceae)**. Molecular Phylogenetics and Evolution 64 (2012) 654–670.

Whittall, JB & Hodges, SA (2007) **Pollinator shifts drive increasingly long nectar spurs in columbine flowers**. Nature, Volume 44, pages 706-709.

Wright, S, Keeling, J & Gillman , L (2006) **The road from Santa Rosalia: A faster tempo of evolution in tropical climates**. Proceedings of the National Academy of the United States of America, Vol. 103, no. 20, pages 7718-7722.

Xiang, Q & Thomas, DT (2008) **Tracking character evolution and biogeographic history through time in Cornaceae—Does choice of methods matter?** Journal of Systematics and Evolution 46 (3): 349–374.

Appendix A - Herbaria accession numbers

Accession numbers of all herbaria specimen used for this thesis. (BOL = Bolus Herbarium, Cape Town, South Africa, NBG = Compton Herbarium, Kirstenbosch National Botanical Garden, Cape Town, South Africa , STEU = Stellenbosch University Herbarium, Stellenbosch, South Africa, WAG = Herbarium Vadense, Wageningen, The Netherlands.)

P. abrotanifolium
WAG
Goldblatt 3940, Herman 409, Schlechter 8265, Goldblatt 4915

P. acetosum
BOL
Paterson 2539, H. Bolus 425
NBG
Batten H.-PL. 71, Taylor 4146, Stayner s.n., Bayliss 2414, Barker 4988

P. aciculatum
STEU
Marais 318, Ward-Hilhorst s.n., Van der Walt 1024, Van der Walt 1039, Marais 265, Van der Walt s.n.

P. acraeum
BOL
Rogers 632, Leighton 3230
NBG
Dlamini 5121, Ward-Hilhorst 19A, Van Jaarsveld 480 sub STEU 777, Compton 28885, Compton 30657, Van Jaarsveld 480, Van Jaarsveld 5979, Kruger s.n., Rudatis 229, Van Jaarsveld 16904

P. aestivale
BOL
Tyson 328, Watermeyer 20058, Gill 65, Bolus 1800, Bolus 13774
NBG
Oliver 5471
STEU
Weber s.n., Lavranos 20952, Marais 146, Van der Walt 1454

P. album
WAG
Van der Walt + Vorster 1344 sub STEU 3086

P. alchemilloides
WAG
Westphal et Westphal-Stevels 1046, Makwarela 64, Schlechter 3979, Bayliss 8090, Vorster 2240, Vorster 2381, Vorster 2372

P. alpinum
WAG
Schlechter 9997

P. anceps ssp. geniculatum
BOL
Fourcade 5800, Fourcade 4847, Holland 3506, Holland s.n., Holland 3545
NBG
Olivier 3001, Tait 75

P. appendiculatum
NBG
Goldblatt & Porter 13537, Goldblatt & Porter 13182, Goldblatt & Porter 13093, Leipoldt s.n.
STEU
Engelbrecht s.n., Van der Walt 1429, Van der Walt 1430

P. aridicola
BOL
Schlechter 11488
STEU
Albers 081086/107, Fischer 125, Van der Walt 1397, Van der Walt 1405, Williamson 3949, Van Zyl s.n., Weber 4

P. aridum
BOL
Bruyns 3347, Acocks sub. N.B.G. 816/53, Bolus 14093, Gilfillan 5514, Bolus 931
WAG
Goldblatt & Manning 10640

P. aristatum
STEU
Fischer & Co 27a, Fischer & Co 28, Schonken 215, Van der Walt 1442, Marais 126, Marais 127

P. asarifolium
BOL
Schlechter 10256, Esterhuysen 20,022, Esterhuysen 16900, Stokoe s.n., Esterhuysen 22,737, Duthie s.n., Duthie s.n.
STEU
Hugo s.n., Van der Walt 573, Van der Walt 1062, Fischer 343, Marais 183, Marais 258, Marais 262, Marais 293

P. auritum var. auritum
BOL
Duthie 152, Estéhuysen 17489, Salter 5671, Leighton 48, Ryder 19988
STEU
Van der Walt 482, Van der Walt 660, Van der Walt 868, Van der Walt 1029, Marais 41/Fischer 340, Marais 53/Fischer 358, Mostert s.n., Marais 95, Marais 96, Marais 138, Marais 188, Cillie s.n., Marais 314

P. auritum var. carneum
BOL
Fourcade 2380, Fourcade 4899, Fourcade 1541, Fourcade 6513, Fourcade 4512, Fourcade 6305, Fourcade 2348, Book 1215, Bolus 9348, Holland 4081, Leighton sub. N.B.G. 1543/50
STEU

Van der Walt 686, Fischer 362, Laubscher s.n., Lavranos 20926, Marais 151, Marais 152, Marais 155, Marais 156, Marais 159, Marais 160, Cowley s.n., Olivier 1928, Van der Walt s.n., Van der Walt s.n.

P. betulinum
NBG
Jones et al. s.n., Fellingham 1523, Cowling 3070, Boucher 521, Benett, Pekeur, Wolfson, Duncan, MSBP 3602, Taylor 5844
WAG
Werdermann et Oberdieck 284, Werdermann et Oberdieck 231, Vorster 2387, Goldblatt 4146, Vlok 1948

P. boranense
WAG
Friis, Gilbert, Rasmussen & Vollesen 92

P. caledonicum
BOL
Bolus 19176, Esterhuysen 4338
STEU
Van der Walt s.n., Vorster s.n., Fischer 254, Fischer 268, Van der Walt s.n.

P. capillare
WAG
Hugo 948

P. carneum
NBG
Helme 3931
STEU
Van Zyl s.n., Van Zyl s.n., Weber s.n., Albers 2602, Van Zyl s.n., Van Zyl s.n., Marais 153, Marais 154, Marais 157, Marais 147, Fischer 363, Drijfhout s.n., Lavranos & Pehlemann 17448, Van der Walt s.n., Van der Walt s.n., Van der Walt s.n., Lavranos 20887, Coutnik s.n.

P. carnosum
WAG
Bayliss 8089, Vorster 2399

P. caroli-henrici
STEU
Marais 130, Hall 2460, Marais 281, Drijfhout 2708, Weber WW 057

P. caucalifolium ssp. caucalifolium
WAG
Thompson 2191

P. ceratophyllum
BOL
Compton sub. NBG 1736/27, Williamson + Leach 2981a
NBG

Littlewood K.J.W. 13/65, Rawe 309/69, Hall 2138 N.B.G. 983/60
WAG
Giess 10824

P. citronellum
WAG
Van Jaarsveld 9593

P. confertum
BOL
Pearson 5703, Pearson 5541
STEU
Van der Walt 956, Van Jaarsveld 4283, Marais 72, Van Zyl s.n., Van Zyl s.n., Lavranos 28338

P. connivens
STEU
Lavranos & Pehlemann 19001, Lavranos & Pehlemann 19000, Lavranos 29902

P. cordifolium
BOL
Bolus 11226, Levyns 10477, Fourcade 4524, Dix 42, Flanagan 2181, Gillett 4576, Gillett 2070, Fourcade 1368, Hutchinson 1265, Fourcade 309, George 4777, Bolus 783, Estèrhuysen 4575, Salter 6797, Estèrhuysen 6487, Pillans 9337, Salter 6741, Leighton 2683, Rodin 1136, Steyn 714, Hutchinson 1198, Duthie 1164

P. coronopifolium
WAG
Werdermann et Oberdieck 506, Schlechter 9983

P. cortusifolium
WAG
de Winter & Hardy 7892

P. crithmifolium
WAG
Nicholas 2647, Schlechter 10880

P. denticulatum
BOL
Compton 5457, Taylor 394, Hutchinson 449, Muis 1070, Salter 3118, Compton 3916, Compton 2563, M.R.L. 1779, Estèrhuysen 28880
WAG
McDonald 942, Schlieben & Ellis 12 357, Bayliss 574

P. dipetalum
BOL
Stokoe 1327, Knysma 2018, Duthie 500, Levyns 11417, Estérhuysen 34910, Fourcade 2018, Duthie sub. N.B.G. 1198/15, Bolus 9902, Leighton 503, Estérhuysen 978, Leighton sub. N.B.G. 671/33
STEU
Van der Walt 1525, Marais 173, Marais 170, Fischer 286, Fischer 277, Vorster 2900, Vorster 2852,

Oliver 5717

P. dolomiticum
WAG
Merxmüller & Giess 28021, Seydel 4010

P. echinatum
BOL
Marloth 4388, Schlechter 8244, Schlechter 8293, Scully 75, Bolus 445, Morris 5605, Morris 5602, Germishuizen 4650, Hardy 587, M.R.L. 6981, Schlechter 122, Leighton sub. N.B.G. 1240/50, Estérhuysen 5733, Estérhuysen 1392, Pillans 5164, Hutchinson 794, Compton 5565, Pillans 5479, Pillans 4950
WAG
Schlechter 8244, Germishuizen 4511, Nicholas 2680, Goldblatt 23

P. elandsmontanum
NBG
Manning 3210
STEU
Marais 449, Marais 450

P. elegans
WAG
Schlechter 9512

P. ellaphieae
BOL
\Bolus 8054, Salter 6525, Bolus 7959, Estérhuysen 35316, Salter 8706
STEU
Marais 197, Marais 204A, Marais 193, Van der Walt 1411, Hudd s.n., Marais 383, Marais 305, Marais 306, Van der Walt 1520

P. elongatum
WAG
Vorster 2413, Williams 861

P. endlicherianum
WAG
Kotschy 1853/90, Von Heldreich s.n.

P. exstipulatum
WAG
Rycroft 3040

P. fasciculaceum
STEU
Van der Walt 1046, Walters 1, Marais 184, Marais 199, Marais 266, Von Willert s.n., Marais 325, Marais 384

P. fergusoniae

BOL
Ferguson 20512, Esterhuysen 19171, Esterhuysen 19562, Esterhuysen 10933, Ryder 19986, Salter 6180
NBG
Bayer 2513, Salter (6185) N.B.G. 1743/36, Schelpe N.B.G. 669/57, Bayer 1711, Bayer 5647, Bayer 2339
STEU
Fischer 231, Fischer 287, Bayer 2513, Meve 281186/396

P. fissifolium
STEU
Van der Walt s.n., Van der Walt s.n., Van der Walt s.n., Fischer 133, Van der Walt 1057, Drijfhout 2657, Marais 108, Marais 142, Marais 219, Marais 227, Marais 228, Marais 231, Marais 232, Marais 233, Marais 234, Marais 235, Marais 236, Marais 238, Marais 239, Marais 244, Marais 248, Marais 249, Marais 250, Marais 251, Marais 348, Marais 354, Lavranos s.n., Albers 2595G, Albers 2595B, Burger s.n., Weber s.n.

P. flavidum
BOL
Banker 20594, Williamson 3094
NBG
Bayliss 3663A, Walters 2319
STEU
Weber s.n.

P. fulgidum
NBG
Helme 4906, Jones et al. 74
WAG
Germishuizen 4798, Schlechter 8057, Goldblatt 2330

P. fumariifolium
BOL
Scully 210, Compton s.n.
NBG
Taylor 11998, Barker 9646, Compton 16386, Compton 13930
STEU
Marais 391, Marais 240

P. fruticosum
BOL
Estérhuysen 19310, Estérhuysen 33899, Estérhuysen 28801, Estérhuysen 28, Muis 936, Hafström 1983 B., Estérhuysen 25926, Estérhuysen 25983, Hafström 1983 c., Estérhuysen 13973, Hutchinson 1158, Estérhuysen 6570, Estérhuysen 6925, George 1612, Estérhuysen 7113, Compton 4233, Leighton 6570, Compton 4028, Trinder-Smith 48, Linder 3214
WAG
Vorster 2367

P. githagineum
BOL

Leighton 254
STEU
Marais 243, Lavranos 20785a

P. glabriphyllum
NBG
Goldblatt 6535A
STEU
Marais 128

P. glechomoides
WAG
Friis 9090, De Wilde 7187, De Wilde 5815, De Wilde & Amshoff 6527, De Wilde 4461, Bamps 6648

P. glutinosum
BOL
Estérhuysen 14009, Estérhuysen 6319, Estérhuysen 26595a, Fourcade 6088, Bremer 296, Pillans 7290, Estérhuysen 26595, George 3840, Hutchinson 1172, Estérhuysen 33855, Page 15644, Kensit sub. N.B.G. 335/14, Estérhuysen 28160

P. grandicalcaratum
WAG
Giess & Müller 14367

P. graveolens
WAG
Meyer 8789, Quintus s.n., Bayliss 7710, Bayliss 7170, Vorster 229

P. grenvillae
BOL
Schlechter 11365, Taylor 1128
STEU
Marais 131, Williamson 3950, Le Roux s.n., Van der Walt 1406, Williamson 3951

P. grossularioides
NBG
Stobie 16, Pretorius 604, Jones et al. 19, Jones et al. 59

P. hermaniifolium
BOL
M.R.L. 4325, Bolus 5134, Estérhuysen 18940, M.R.L. 4396, M.R.L. 6255, Estérhuysen 4287, Estérhuysen 970, Estérhuysen 18773
WAG
Schlechter 9767

P. hirtum
WAG
Bos 424

P. hispidum
BOL
Estérhuysen 4032, Wilman 13545, Estérhuysen 16640, Estérhuysen 22331, Estérhuysen 25633, Estérhuysen 29954, Stokoe 9044, Bolus 5324, Bolus 2601, Compton 10141, Wilman 14099, Bolus 550, Stokoe 2653, Schlechter 9848, Pillans 9309, Pillans 9469, Estérhuysen 672
WAG
Schlechter 9848

P. hypoleucum
WAG
Schlechter 9758

P. hystrix
BOL
Leighton 258, Compton 2918, Archer 132
NBG
Compton 13926, Compton 12089, Compton 12158, Compton 12057, Barker 6791, Comptom 21182

P. incrassatum
BOL
Bolus 19345, Stokoe 8456, Estérhuysen 5699, Bolus 1112, M.R.L. 6985, M.R.L. 4031, Williamson 3106, Schlechter 11005, Compton 5500, Taylor 1061, Bean & Viviers 2569, Leroux 2693
STEU
Craib s.n., Craib s.n., Weber 033, E.M. Marais 60, Drijfhout 2971A, E.M. Marais 284, E.M. Marais 285, E.M. Marais 132, E.M. Marais 74, J.J.A. Van der Walt 1398, Drijfhout 2783, Boucher 63, Drijfhout 2942A, Van der Walt 789, Fischer 1, M. Schonken 165, D.A. Boucher 73, D.A. Boucher s.n.

P. inquinans
WAG
Bos 53, Bayliss 7511A, Bayliss 6244, Olivier 2147

P. ionidiflorum
WAG
Vorster 2326

P. karooicum
WAG
Goldblatt 6544

P. ladysmithianum
NBG
Vlok 2027
STEU
Lavranos & Pehlemann s.n., Lavranos & Pehlemann 17535a, Vlok 2557

P. laevigatum ssp. laevigatum
BOL
Bolus 2276, Compton 3917, Acocks 19917, Fourcade 758, Esterhuysen 278, Rodit 1109, Häfstrom & Acock 753, Esterhuysen 6661, Compton 10525, Esterhuysen 7040, Esterhuysen s.n., Compton

4516, Bolus 11739, Esterhuysen s.n., Esterhuysen 22,797, Esterhuysen 4689, Gill 19, Fourcade 3217, Fourcade 2704, Fourcade 2091
NBG
Thompson 2195, Taylor 9796, Boshoff 360, Oliver 5330, Van Wyk 384, Boucher 38, Oliver 5566, Oliver 4535, Hugo 1451, Geldenhuys 467
WAG
Bayliss 6382

P. lanceolatum
WAG
Hugo 2289

P. leipoldtii
STEU
Marais 317, Marais 222, Marais 220, Marais 216, Marais 207, Marais 215, Marais 109, Marais 98, Marais 100, Walters 2, Drijfhout 2821, Lavranos & Pehlemann 17480, Van Zyl s.n., Weber s.n., Weber s.n., Marais 221, Van Niekerk s.n., Muller 4038, Muller 4036, Marais 102

P. leptum
BOL
Leipoldt 19185
NBG
Oliver 5797
STEU
Oliver 4981, Marais s.n., Marais 209

P. longicaule var. longicaule
BOL
Fourcade 4910, Werdermann + Oberdieck 343, Page s.n., Page s.n., Leighton 3059, Burman 938, Esterhuysen 24,351, Gillet 4118, Pillans 8049, Pearson 5220, Pillans 8765, Esterhuysen 23,767, Esterhuysen s.n., Leighton 215777, Esterhuysen 22,089, Salter 6478, Pillans 5197, Bolus s.n., Leighton 3059, Salter 6442, Hutchinson 18, Esterhuysen s.n., Esterhuysen 341, Stokoe 7398, Leighton s.n., Leighton 3149, Barker 5953
NBG
Germishuizen 4195, Barker 5953, Salter 7004, Barker 4191, Compton 20106, Barker 1695, Barker 8132, Barker 3183, Compton 11647, Barker 5953, Van Wyk 2007, Van Jaarsveld 5724, Wilmot 630/76, Van Niekerk 456, Compton 23645, Compton 24213, Taylor 4021, Guthrie 279, Compton 18205, Barker 1691, Barker 1148, Wasserfall 526, Jamieson 47, Ward-Hilhorst 16A, Steiner 3045, Viviers 736, Kruger 295, Drijfhout 4116, De Kock 10, Oliver 5499, Orchard 288, Thompson 1451, Durand 92, Osrin 3, Kruger 764, Lamb 104, Low 795, Boucher 2254, Jardine & Jardine 264, Hanekom 979, Boucher 3277, Taylor 9632, Goldblatt & Manning 10382, McDonald 808, Hugo 165, Hugo 750, Smith 13, Van Zyl 3461, Emdon 49

P. longiflorum
BOL
Leipoldt 3079, Leipoldt 4093, Salter 6462, Adamson 1515
NBG
Forrester s.n., Walter 88, Leipoldt 4093, Le Roux 2828, Jordaan 122
STEU
P. Drijfhout 1813, C.M. Schonken 37, E.M. Marais 35 / Fischer 333, I.S. Walters 3, J.J.A. van der

Walt 1412, J.J.A. van der Walt 1421, E.M. Marais 182, E.M. Marais 267, E.M. Marais 268, E.M. Marais 308, Le Maitre s.n., Struck 50153

P. luridum
WAG
Werdermann & Oberdieck 1412, Torre & Correia 13.115, Bamps, Symoens & Van den Berghen 217, Macuácua 1265, Phillips 3777 B, Werdermann & Oberdieck 1409, Lovett & Kayombo 4856A, Goldblatt, Brummitt & Lovett 8224, Prins-Lampert 457, Jordaan 3294, Werdermann & Oberdieck 1231, Torre 3344, Pawek 10227, Kemp 1121, Stolz 100, Balsinhas 03061, Germishuizen 2889, Werdermann & Oberdieck 1511

P. luteum
STEU
Fischer 33, Marais 120, Perry 3243, Marais 270, Marais 271, Marais 121

P. magenteum
BOL
Levyns 11,217, M.R.L. 1803, Koutnik 1199, M.R.L. 1513, Bean & Viviers 1743, Schlechter 8662, Compton sub. N.B.G. 351/25, Hutchinson 428, Estérhuysen 33963, Estérhuysen 20,589a, Estérhuysen 5315, Leipoldt 3086, Barker 20698, Leipoldt 3800, Estérhuysen 5771, Hafström & Acock 742, Schlechter 5026, Adamson 739
WAG
Schlechter 8662, Goldblatt 4067, Goldblatt 11407, Taylor & Midgley 19

P. minimum
WAG
Giess 14675

P. moniliforme
BOL
Hutchinson 827, Leipoldt 4394, Pillans 14143, Compton 3306, Kolbe 14293, Leighton 3186, Marloth 12489, Leighton 1124, Hardy 80
NBG
Helme 5802, Jardine & Jardine 558, Goldblatt & Porter 12782
STEU
Weber s.n., Bruyns 1516, Van der Walt s.n., Lavranos & Pehlemann 17478, I.S. Walters 5, E.M. Marais 71, D.A. Boucher 77, E.M. Marais 125, E.M. Marais 76, E.M. Marais 75, E.M. Marais 73, E.M. Marais 210, E.M. Marais 217, E.M. Marais 218, E.M. Marais 224, E.M. Marais 68, E.M. Marais 119, E.M. Marais 123, E.M. Marais 133, E.M. Marais 252, E.M. Marais 272, E.M. Marais 282, E.M. Marais 225, E.M. Marais 273, E.M. Marais 275, E.M. Marais 335, E.M. Marais 355, Van Zyl s.n.

P. multibracteatum
WAG
Bos 8227, Jansen 7148, Jansen 6511, Jansen, de Wit & Aneke 4578, Jansen 4472, Jansen 4263, Jansen 4041, Jansen 3788, Wieringa 4982, Wieringa 4918, Wieringa 5008, Jansen 1559, Jansen 3454, Jansen 2947, Jansen 2061, Jansen 1982, De Wilde 6631, De Wilde 6642, De Wilde 4720, De Wilde 4497, Friis, Bidgood, Host, Wondafrash & Kebede 6655, de Wilde et de Wilde-Duyfjes 9640, de Wilde et de Wilde-Duyfjes 8728, de Wilde et de Wilde-Duyfjes 8654, de Wilde et de Wilde-Duyfjes 6059, Chojnacki 94, Westphal et Westphal-Stevels 1046, Westphal et Westphal-Stevels

2382, Westphal et Westphal-Stevels 1503, Gilbert & Thulin & Aweke 261, de Wilde et de Wilde-Duyfjes 10883, de Wilde et de Wilde-Duyfjes 10531, de Wilde et de Wilde-Duyfjes 6631, Lucas & Williams EA 12326

P. myrrhifolium var. myrrhifolium
BOL
Estérhuysen 32944, Estérhuysen 26470, Pillans 8635, M.R.L. 3623, Bolus 8419, Bolus 4254, Estérhuysen 24,350, Pillans 9584, Compton 16, Scully 168, M.R.L. 7353, M.R.L. 1415, M.R.L. 1171, M.R.L. 3168, M.R.L. 1427, Pillans 9653

P. nanum
WAG
Schlechter 9991, Pearson 2837, Schlechter 11105

P. nervifolium
BOL
Leighton 3184, Esterhuysen 23256, Leighton s.n., Leipoldt 3078, Compton 3811, Leipoldt s.n.
STEU
Marais 334, Marais 274, Marais 276, Van der Walt s.n., Lavranos & Pehlemann 18999, Marais 141, Marais 145, Marais 253

P. oblongatum
BOL
Herre 2891, Scully 64, Herre 2947
STEU
Van der Walt s.n., Van Jaarsveld 5368, Williamson 4463, Weber s.n., Jones s.n.

P. odoratissimum
BOL
J.R. & B.R. 125, J.R. & B.R. 407, Fourcade 2707, Fourcade 2749, Fourcade 5140, Fourcade 3072, Muir 1801, Leighton sub. N.B.G. 671/32, Estérhuysen 4340
WAG
Fries, Norlindh et Weimarck 548, Retief 342, Olivier 2131

P. ovale
WAG
Fries, Norlindh et Weimarck 1641, Schlechter 9137, Schlechter 9804, Hugo 607, Hugo 921

P. pallidoflavum
STEU
Marais 208, Marais 180, Marais 192, Marais 201, Marais 202, Marais 303, Marais 311, Marais 323

P. paniculatum
WAG
Giess & Müller 14348

P. papilonaceum
WAG
Drewe 630, Williams 846

P. parvipetalum
BOL
Leipoldt 20760
STEU
Bruyns 1519, Le Roux s.n., Stirton 9242, Marais 327, Oliver 9855

P. patulum
WAG
de Hoogh 14, Vorster 2415, Schlechter 9993, Schlechter 9856, Schlechter 9191

P. peltatum
WAG
Vorster 2254, Olivier 2234, Olivier 1990, Bayliss BRI B 123

P. petroselenifolium
BOL
Leipoldt 4003
NBG
Forrester 499, Perry 3059
STEU
Lavranos & Pehlemann 17417, Marais 51/Fischer 356, Marais 65, Van der Walt 1625

P. pilosellifolium
STEU
Schonken 201, Schonken 213, Van der Walt 1105, Fischer 293, Marais 18/Fischer 313, Marais 34/Fischer 332, Lavranos & Pehlemann 18901, Fischer 364, Fischer 365, Lavranos 20911, Thomas 79, Marais 263, Marais 264, Marais 279, Marais 280, Stirton 11505, Marais 307B, Marais 344, Van Zyl s.n., Marais 398, Weber s.n.

P. pinnatum
BOL
Muir 1268, Salter 4233, Bolus 3068, Estérhuysen 22,394, Salter 2909, Salter 7903, Salter 1908, Salter 2940, Salter 5721, Estérhuysen 20,836, Estérhuysen 25492, Leighton 1571, Leighton 2560, Bolus 19181, Hafström & Acock 1976, Estérhuysen 21,073, Estérhuysen 4312, Bolus 20513
NBG
Boucher 1438, Boucher 961, Rode & Boucher 0202
STEU
Fischer 271, Marais 32/Fischer 330, Marais s.n./Fischer 338, Van der Walt 510, Drijfhout 1625, Van der Walt 662, Hugo s.n., Van der Walt s.n., Vorster 2905, Vorster 2917, Van der Walt 1101, Lavranos 20905, Marais 82, Muller s.n., Marais 169, Marais 289, Van der Walt 1558, Van der Walt 1562, Van der Walt 1572, Van der Walt 1573, Marais 340, Marais 346, Buys 88, Marais 405, Marais 385, Marais 381, Marais 407, Buys s.n., Van der Merwe s.n.

P. praemorsum
WAG
Schlechter 11006

P. proliferum
STEU
Van der Walt 483, Van der Walt 546, Van der Walt 651, Van der Walt 657, Van der Walt 931, Van der

Walt 1083, Van der Walt 1084, Van der Walt 1086, Van der Walt 1093, Fischer 285, Fischer 302, Fischer 311, Marais 44, Vorster 2925, Van der Walt 1425, Van Zyl s.n., Buys s.n., Marais 369, Marais 373, Marais 375, Van der Walt s.n., Richfield s.n., Weber s.n.

P. pulchellum
WAG
Evrard 8865, Schlechter 11008, Goldblatt 2793

P. punctatum
BOL
Pillans s.n.
STEU
Fischer 34, Van der Walt 944, Lavranos & Pehlemann 18876, Marais 67, Marais 69, Meve 273, Weber 2, Weber s.n.

P. quarciticola
NBG
Schmiedel 109738, Nordenstam & Lundgren 1415, Helme 4908

P. quercifolium
BOL
Fourcade 117, Fourcade 3613, Levyns 10562, Estérhuysen 24,963, Estérhuysen 6392, Estérhuysen s.n., Ryder 17, Bolus 304, Bolus 14499, Acocks 19909, Niekerk 468, Estérhuysen 412, Estérhuysen 4699, Hops 2, West 219, Fourcade 4631, Fourcade 5686
WAG
Vorster 2362, Bayliss 7707, Gien 1564, Bayliss 6564, Coppejans EC 1377, Bayliss 6032

P. quinquelobatum
WAG
Bos 8104, Westphal et Westphal-Stevels 1669, Westphal et Westphal-Stevels 2795, De Wilde 6893, De Wilde & Ebba 5039, De Wilde 6978, De Wilde 5110, Friis, Bidgood, Host, Wondafrash & Kebede 6660, Jansen 7003, Jansen 3617, Jansen 3455, Pedersen 811, Faden and Evans 74/585, Lovett 3177, Williams EA 12324, Beentje 1791, Bidgood, Mwasumbi and Vollesen 1124

P. radens
WAG
Olivier 2158

P. radiatum
STEU
Van der Walt 1063, Van der Walt 1494, Marais 386, Albers s.n., Van der Walt 970

P. radulifolium
BOL
Holland 3569, Muir 2971, Urton sub. N.B.G. 79/55, Esterhuysen 14098, Bolus 11738, Bolus 2273, Tugwell sub. N.B.G. 2649/57, Fourcade 2403, Fourcade 1784, Fourcade 3794, Fourcade 3394, Fourcade 3486, Fourcade 2103

P. rapaceum
STEU

Lavranos & Beeck 20915, Brits s.n., Drijfhout 1370, Drijfhout 1300, Drijfhout 1432, Van der Walt s.n., Van der Walt s.n., Van der Walt 575, Boucher 49, Marais s.n., Boucher s.n., Van der Walt 661, Van der Walt s.n., Van der Walt 750, Schonken 36, Fischer 3, Fischer 14, Fischer 35, Van der Walt 816, Schonken 87, Ward 44A, Van Jaarsveld 4271, Van der Walt 1026, Van der Walt 1035, Van der Walt 1043, Van der Walt 1048, Fischer 238, Fischer 303, Fischer 310, Marais 25/Fischer 322, Fischer 329, Lavranos & Pehlemann 18846, Drijfhout 2943, Lavranos & Pehlemann 19742, Lavranos & Pehlemann 19832, Stirton 10050, Lavranos 20889, Lavranos 20916, Le Roux s.n., Marais 97, Marais 122, Cillie s.n., Maggs 44, Williamson 3528, Marais 200, Marais 283, Marais 322, Marais 378, Weber 1, Marais 388, Marais 393, Marais 402, Marais 408

P. reflexipetalum
BOL
Esterhuysen 3382, Esterhuysen 21,936, Leighon 3158, Bolus 8943
STEU
Van der Walt s.n., Fischer 112, Marais 185, Marais 203, Marais 205, Marais 302, Marais 397

P. reflexum
NBG
Perry 1987, Snyman & Manning 1526
STEU
Van Wyk 161, Marais 278, Lavranos 29880, Craib s.n.

P. reniforme
WAG
Bayliss 5658, Olivier 2163, Vorster 2316, Vorster 2324

P. scabrum
WAG
Goldblatt 4162, Grant s.n., Edwards 150, Coppejans 1496, Vorster 2403, Schlechter 8513, Schlechter 7736, Goldblatt 6531, Werdermann et Oberdieck 523

P. senecioides
WAG
Bayliss 6112, Merxmüller & Giess 28793, Goldblatt 3213, Goldblatt 3036, Goldblatt 4213, Vorster 2393, Goldblatt 5663, Seydel 4409, Coppejans 1469

P. setulosum
WAG
Schlechter 9805

P. spinosum
WAG
Merxmüller & Giess 28 629

P. sublignosum
BOL
Esterhuysen 18430, Esterhuysen 14688, Smuts & Gillet 3479, Esterhuysen 15339, Esterhuysen 22,513, Esterhuysen 28416, Esterhuysen 21,906, Schlechter 9976, Esterhuysen 22,513, Esterhuysen 25701

P. suburbanum ssp. bipinnatifidum
BOL
Pillans 2699, Pillans 2736, Pillans 3312, Hafström 1979, Salter 8275, Esterhuysen 4322, Pillans 8201, Leighton 1651, Esterhuysen 4332
NBG
Hanekom & Walsh 176

P. suburbanum ssp. suburbanum
BOL
Levyns 1479, Rogers 2978, Holland 3725, Leighton s.n., Fourcade 1969a, Fourcade 3300, Fourcade 1857
NBG
Kruger 1214, C.R.E.W. CR42

P. tenuicaule
BOL
Hall s.n., Bolus 6658
NBG
Le Roux 4539, Hugo 2805, Hall 787, Oliver, Tölken & Vennter 472, Van der Westhuizen 131/80, Thompson & Le Roux 376, Van Jaarsveld 6223, Van Jaarsveld 4309A, Marloth 12341, Lavranos, Pehlemann & Barad 19222 sub STEU 2955, Viviers 2063, Van Jaarsveld 4309, Jamieson s.n., Hall 1023, Hall 818, Van Jaarsveld & Kritzinger 6223

P. ternatum
WAG
Schlechter 5565, Werdermann et Oberdieck 870, Bos 698

P. ternifolium
BOL
Acock s.n., Acock s.n., Esterhuysen s.n., Esterhuysen 20,993, Salter 6566, Esterhuysen 21,195, Pillans 9994, Esterhuysen 15727, Esterhuysen s.n., Duthie 1075, Duthie 1075a, Duthie s.n.
STEU
Duthie s.n., Garside s.n., Duthie s.n., Marais s.n., Drijfhout 1535, Marais 382, Marais 337, Marais 319, Marais 320, Marais 164, Van der Walt s.n., Drijfhout 1627, Drijfhout 262

P. tetragonum
WAG
Olivier 1702

P. tomentosum
BOL
Acock s.n., Acock s.n., Esterhuysen s.n., Esterhuysen 20,993, Salter 6566, Esterhuysen 21,195, Pillans 9994, Esterhuysen 15727, Esterhuysen s.n., Duthie 1075, Duthie 1075a, Duthie s.n.
WAG
Vlok 2052

P. tragacanthoides
BOL
Acocks 5538, Bolus 1784, Manoergh 17529, Esterhuysen 2719, Esterhuysen s.n., Pearson 5906, Esterhuysen 19739, Flanagan 1542

P. triandrum
BOL
Leighton 3346
STEU
Van der Walt & Vorster 1276, Craib s.n., Van der Walt s.n., Van der Walt s.n., Friedrich 452, Van der Walt 1278, Van Niekerk s.n., Van Zyl s.n.

P. tricolor
WAG
Marshall 174

P. triphyllym
BOL
Esterhuysen 4311, Esterhuysen s.n., Esterhuysen 18020, Esterhuysen 18111, Leipoldt 4005, Leipoldt 4006, Esterhuysen 19775, Esterhuysen 22451
NBG
Taylor 11953, Compton 16769, Compton 16772, Bayer 3172, Barker 1299, Compton 16773
STEU
Cillie s.n., Von Willert s.n.

P. triste
WAG
Bos 462, Bos 423, Bos 422, Bayliss 6315, Williams 350, Vorster 2417, Bos 508

P. undulatum
BOL
Levyns 1002, Levyns 2436, Lewis sub N.B.G. 2789/32, Hall sub N.B.G. 752/50, Leipoldt s.n., Barker & Lewis 20601
STEU
Van der Walt s.n., Lavranos & Pehlemann 17470, Van der Walt 1111, Lavranos & Pehlemann 18803, Muller 4041b, Marais 331, Van der Walt 1593

P. vinaceum
STEU
Williamson 4445, Marais 77, Williamson 3527, Van der Walt & Vorster 1275, Lavranos 20785, Venter 8630, Van Jaarsveld 19.54*, Visser s.n., Williamson 4010, Von Willert s.n., Williamson 4341, Weber s.n., Williamson & Hammer 4465

P. violiflorum
BOL
Leighton 21158, Barker 1301
NBG
Marloth 11824, Walters 2712, Barker 1301, Murtry N.B.G. 405/67
STEU
Fischer 217, Fischer 216

P. vitifolium
BOL
Esterhuysen 23,545, Leighton 3457, Leighton 4168, Esterhuysen 18330, Salter 6392, Page 14201,

Esterhuysen 30839, Martley s.n., Bolus 8527, Schlechter 9253, Duthie 1161
NBG
Powrie 291, Bolus s.n., Vorster & Van der Walt 2929 sub STEU 2916, Duthie 1161, Fourcade 3301

P. whytei
WAG
Ash 2536, Gillett 18325, Croockewit 205

P. wuppertalense
STEU
Van der Walt 750, Fischer 116, Fischer 118, Fischer 119, Van der Walt 1044, Marais 114, Marais 115, Marais 116, Marais 117, Marais 401, Bruyns 1518, Lavranos & Pehlemann 17482A

P. zonale
WAG
Lingér 44, Vorster 2348, Grant s.n., Vorster 2354, Bayliss BRI.B. 25

Appendix B - Median nectar spur lengths

Species	Median (in cm)	Species	Median (in cm)	Species	Median (in cm)
P. abrotanifolium	2.6	P. exstipulatum	1.1	P. pallidoflavum	4.2
P. acetosum	2.5	P. fasciculaceum	5.2	P. paniculatum	0.2
P. aciculatum	1.5	P. fergusoniae	2.3	P. papilionaceum	0.4
P. acraeum	2.7	P. fissifolium	3.7	P. parvipetalum	1.1
P. aestivale	4.8	P. flabellifolium	3.8	P. patulum	0.6
P. album	1.2	P. flavidum	2.8	P. peltatum	3.0
P. alchemilloides	2.8	P. frutetorum	3.8	P. petroselenifolium	1.8
P. alpinum	3.4	P. fruticosum	1.5	P. pillosellifolium	1.2
P. alternans	0.6	P. fulgidum	2.7	P. pinnatum	2.0
P. althaeoides	0.3	P. fumarifolium	2.4	P. praemorsum	2.6
P. anceps	0.4	P. geniculatum	0.4	P. proliferum	1.0
P. anethifolium	2.8	P. gibbosum	2.4	P. pseufumarioides	0.3
P. antidysentericum	2.0	P. githagineum	1.1	P. pulchellum	4.3
P. appendiculatum	8.4	P. glabriphyllum	2.1	P. pulverulentum	3.5
P. aridicola	2.7	P. glechomoides	1.5	P. punctatum	2.6
P. aridum	4.4	P. glutinosum	0.7	P. quarciticolae	1.0
P. aristatum	2.7	P. grandicalcaratum	1.2	P. quercetorum	1.8
P. articulatum	6.0	P. graveolens	0.6	P. quercifolium	0.9
P. asarifolium	0.7	P. grenvilleae	2.6	P. quinquelobatum	3.2
P. auritum ss. auritum	1.3	P. griseum	0.6	P. radens	1.0
P. auritum ss. carneum	1.3	P. grossularioides	0.2	P. radiatum	4.7
P. australe	0.3	P. havlase	0.3	P. radulifolium	3.7
P. barklyi	4.8	P. hermanniifolium	0.8	P. rapaceum	1.8
P. betulinum	0.7	P. hirtum	0.3	P. reflexipetalum	1.2
P. boranense	2.5	P. hispidum	0.3	P. reflexum	1.7
P. bowkeri	4.3	P. hypoleucum	0.9	P. reniforme	2.6
P. buysii	0.1	P. hystrix	3.7	P. rotundipetalum	0.3
P. caffrum	2.5	P. incarnatum	1.3	P. scabrum	0.5
P. caledonicum	0.9	P. incrassatum	3.6	P. schizopetalum	6.8
P. capillare	1.6	P. inquinans	3.0	P. senecioides	0.3
P. carneum	5.0	P. iocastum	0.7	P. sericifolium	4.8
P. carnosum	0.4	P. ionidiflorum	3.2	P. setulosum	0.9
P. carolihenrici	2.6	P. karooicum	1.2	P. spinosum	1.7
P. caucalifolium	3.2	P. ladysmithianum	3.5	P. stipulaceum	5.0
P. caylae	1.8	P. laevigatum	2.2	P. sublignosum	0.9
P. ceratophyllum	0.8	P. lanceolatum	2.6	P. suburbanum ssp. bipinnatifidum	2.8
P. citronellum	0.6	P. laxum	0.6	P. suburbanum ssp. suburbanum	1.7
P. confertum	1.2	P. leipoldtii	1.1	P. tenuicaule	1.4
P. connivens	4.9	P. leptum	1.9	P. ternatum	0.9
P. cordifolium	0.7	P. leucophyllum	0.9	P. ternifolium	0.8
P. coronopifolium	0.6	P. lobatum	2.8	P. tetragonum	3.4
P. cotyledonis	0.3	P. longicaule	2.2	P. tomentosum	1.7
P. crassicaule	2.0	P. longiflorum	2.7	P. tongaense	3.5
P. crithmifolium	0.5	P. luridum	4.6	P. torulosum	5.0
P. cucculatum	1.0	P. luteolum	1.9	P. traganthoides	0.7
P. dasyphyllum	0.7	P. luteum	2.7	P. triandrum	2.8
P. denticulatum	0.4	P. magenteum	3.6	P. tricolor	0.2
P. desertorum	1.3	P. mollicomum	1.8	P. trifidum	2.0
P. dichondrifolium	2.9	P. moniliforme	3.3	P. triphyllum	1.2
P. dipetalum	1.1	P. multibracteatum	4.1	P. triste	3.7
P. dolomiticum	0.7	P. multiradiatum	3.0	P. undulatum	0.9
P. drummondii	0.3	P. mutans	1.5	P. vinaceum	1.8
P. echinatum	3.1	P. myrrhifolium	0.7	P. violifloreum	0.9
P. elandsmontanum	1.0	P. nanum	0.2	P. vitifolium	0.5
P. elegans	1.4	P. nervifolium	3.5	P. whytei	0.8
P. ellaphieae	1.3	P. oblongatum	6.0	P. worcesterae	1.0
P. elongatum	2.7	P. odoratissimum	0.8	P. wuppertalense	4.1
P. endlicherianum	3.2	P. otaviense	1.3	P. xerophyton	1.9
P. exhibens	1.5	P. ovale	0.6	P. zonale	2.6

Appendix C - R scripts

In this appendix I have included the most important R scripts I used for the analyses I performed during this project. In most cases they can be copied straight to R, although I recommend using a text editor like Tinn-R. A few important remarks:

- for each script, I assume the libraries APE, Geiger, Phytools, Auteur and Diversitree have been loaded;
- in all scripts, 'phy' refers to a single tree, 'trees' to a block of trees, and 'spurs' to a file with spur lengths and corresponding species names;
- I assume a working directory has been set using the command 'setwd';
- some scripts will contain notes, which are headed by a hashtag (#);
- specific settings of models will need to be adjusted to the data being used;
- all analyses were performed with R version 2.14.2, results may vary with different versions.

Auteur

```
# for this analysis a file with standard deviations of spur length is needed ('spurs.sd'). Settings like number of frequencies, sample frequency, et cetera, may be adjusted.
# run two chains
r=paste(sample(letters,9,replace=TRUE),collapse="")
lapply(1:2, function(x) rjmcmc.bm(phy=phy, dat=spurs, SE = spurs.sd, ngen=1000000, sample.freq=100, prob.mergesplit=0.2, simplestart=FALSE,
prop.width=NULL, fileBase=paste(r,x,sep=".")))

# collect directories
dirs=dir("./",pattern=paste("BM",r,sep="."))
pool.rjmcmcsamples(base.dirs=dirs, lab=r)

## view contents of .rda
load(paste(paste(r,"combined.rjmcmc",sep="."),paste(r,"posteriorsamples.rda",sep="."),sep="/"))
print(head(posteriorsamples$rates))
print(head(posteriorsamples$shifts))

## plot Markov sampled rates
dev.new()
shifts.plot(phy=phy, base.dir=paste(r,"combined.rjmcmc",sep="."), burnin=0.5, legend=TRUE, edge.width=2)


# To assess the effective sample size (ESS): copy the combined log file from the directory and paste it to the main workspace. Read it with (change 'zsqtuyirm' to the the random letter combination that has been created in the previous step)
x <- read.table(file="zsqtuyirm.rjmcmc.log", header=T)
# For some reason the sixth(?) column ("root") generates NA values, so this has to be removed by using
x_subset <- x[c(1,2,3,4,5,7,8,9)]
# Now, effective sample sizes can be caluclated for each column using coda:
#library(coda)
effectiveSize(x_subset)
```

```
# clean-up: unlink those directories
unlink(dir(pattern=paste(r)),recursive=TRUE)
```

ACE

```
# To find the oldest MRCA (=Most Recent Common Ancestor) of two taxa:
oldest.mrca <- function(tree, tips) {
H <- nodeHeights(tree)
X <- mrca(tree)
n <- length(tips)
nodes <- height <- vector(); k <- 1
for (i in 1:(n-1)) for (j in (i+1):n) {
nodes[k] <- X[tips[i], tips[j]]
height[k] <- H[match(nodes[k], tree$edge[,1]),1]
k <- k+1
}
z <- match(min(height), height)
return(nodes[z])
}

MRCA <- list()
for (i in 1:length(trees)) {
MRCA[[i]] <- oldest.mrca(trees[[i]],c("species1","species2", et cetera) }
nodeNumber <- MRCA

node <- list()
for (i in 1:length(trees))  {
quotationMarks <- function(xvar) deparse(xvar)
node[[i]] <- quotationMarks(nodeNumber[[i]])  }


# Check whether species names of spur data and trees correspond
name.check <- list()
all <- list()
for (i in 1:length(trees)) {
name.check[[i]] <- name.check(trees[[i]], spurs)
all[[i]] <- all(trees[[i]]$tip.label %in% names(spurs))   }
unique(name.check)
unique(all)

# Maximum Likelihood
aceML <- list()
for (i in 1:length(trees))  {
aceML[[i]] <- ace(spurs, trees[[i]], type="continuous", method="ML")  }

# Restricted Maximum Likelihood
aceREML <- list()
for (i in 1:length(trees))  {
aceREML[[i]] <- ace(spurs, trees[[i]], type="continuous", method="REML")  }
```

```
# Squared-Change Parsimony
aceSCP <- list()
treesBRLone <- list()
for (i in 1:length(trees))  {
treesBRLone[[i]] <- compute.brlen(trees[[i]], 1)
aceSCP[[i]] <- ace(spurs, treesBRLone[[i]], type="continuous", method="ML")  }

node343ML <- list()
for (i in 1:length(trees))  {
node343ML[[i]] <- aceML[[i]]$ace[[node[[i]]]]  }
# Results in a list with all reconstructed values for the node of interest

averageML <- (do.call(sum, node343ML))/(length(trees))
# Average value for the node of interest

# Lower value of 95% Credibility Interval (=95% CI)
aceML95lower <- list()
for (i in 1:length(trees))  {
aceML95lower[[i]] <- aceML[[i]]$CI95[,1][[node[[i]]]]  }

# Upper value of 95% CI
aceML95upper <- list()
for (i in 1:length(trees))  {
aceML95upper[[i]] <- aceML[[i]]$CI95[,2][[node[[i]]]]  }

# (May be repeated for results of SCP and REML)

averagelowerML <- (do.call(sum, aceML95lower))/(length(trees))
averageupperML <- (do.call(sum, aceML95upper))/(length(trees))

node12ML <- c(averagelowerML, averageML, averageupperML)
# Results in reconstructed values, plus 95% CI, averaged over all trees, for node of interest
```

fitContinuous

```
# The first part of these scripts have been copied from the online help function of fitContinuous
#---- STORE RESULTS
brownFit <- list()
for (i in 1:length(trees)) {
brownFit[[i]] <-  fitContinuous(trees[[i]], spurs, bounds=list(beta=c(0,10000)))      }

aic.brown <- numeric()
for (i in 1:length(trees)) {
aic.brown[[i]]<-brownFit[[i]]$Trait1$aic  }

#----------------------------------------------------
#   PHYLOGENETIC SIGNAL: FIT LAMBDA
#----------------------------------------------------
```

```
lambdaFit <- list()
for (i in 1:length(trees)) {
lambdaFit[[i]]<-fitContinuous(trees[[i]], spurs, model="lambda", bounds=list(beta=c(0,1000)))   }

# Compare likelihoods:

d.lambda<-numeric()
for(i in 1:length(trees)) d.lambda[i]=2*(lambdaFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.lambda=pchisq(d.lambda, 1, lower.tail=FALSE)

aic.lambda<-numeric()
for(i in 1:length(trees)) aic.lambda[i]<-lambdaFit[[i]]$Trait1$aic

#-----------------------------------------------------
#    TIME PROPORTIONALITY: DELTA
#-----------------------------------------------------

deltaFit <- list()
for (i in 1:length(trees))  {
deltaFit[[i]]<-fitContinuous(trees[[i]], spurs, model="delta", bounds=list(beta=c(0,1000)))  }

# Compare likelihoods:

d.delta<-numeric()
for(i in 1:length(trees)) d.delta[i]=2*(deltaFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.delta=pchisq(d.delta, 1, lower.tail=FALSE)

aic.delta<-numeric()
for(i in 1:length(trees)) aic.delta[i]<-deltaFit[[i]]$Trait1$aic

#-----------------------------------------------------
#   SPECIATIONAL MODEL: KAPPA
#-----------------------------------------------------
kappaFit <- list()
for (i in 1:length(trees)) {
kappaFit[[i]]<-fitContinuous(trees[[i]], spurs, model="kappa", bounds=list(beta=c(0,1000)))  }

# Compare likelihoods:

d.kappa<-numeric()
for(i in 1:length(trees)) d.kappa[i]=2*(kappaFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.kappa=pchisq(d.kappa, 1, lower.tail=FALSE)

aic.kappa<-numeric()
```

```
for(i in 1:length(trees)) aic.kappa[i]<-kappaFit[[i]]$Trait1$aic

#-----------------------------------------------
#   OU MODEL: ALPHA
#-----------------------------------------------
ouFit <- list()
for (i in 1:length(trees)) {
ouFit[[i]]<-fitContinuous(trees[[i]], spurs, model="OU", bounds=list(beta=c(0,1000)))  }

# Compare likelihoods:

d.ou<-numeric()
for(i in 1:length(trees)) d.ou[i]=2*(ouFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.ou=pchisq(d.ou, 1, lower.tail=FALSE)

aic.ou<-numeric()
for(i in 1:length(trees)) aic.ou[i]<-ouFit[[i]]$Trait1$aic

#-----------------------------------------------
#   EARLY BURST MODEL: R
#-----------------------------------------------
ebFit <- list()
for (i in 1:length(trees))  {
ebFit[[i]]<-fitContinuous(trees[[i]], spurs, model="EB", bounds=list(beta=c(0,10000)))  }

# Compare likelihoods:

d.eb<-numeric()
for(i in 1:length(trees)) d.eb[i]=2*(ebFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.eb=pchisq(d.eb, 1, lower.tail=FALSE)

aic.eb<-numeric()
for(i in 1:length(trees)) aic.eb[i]<-ebFit[[i]]$Trait1$aic

#-----------------------------------------------
#   White Model:
#-----------------------------------------------
whiteFit <- list()
for (i in 1:length(trees))  {
whiteFit[[i]]<-fitContinuous(trees[[i]], spurs, model="white", bounds=list(beta=c(0,1000)))  }

# Compare likelihoods:

d.white<-numeric()
for(i in 1:length(trees)) d.white[i]=2*(whiteFit[[i]]$Trait1$lnl-brownFit[[i]]$Trait1$lnl)
```

```
# Calculate p values assuming chi-squared distribution with 1 d.f.
p.white=pchisq(d.white, 1, lower.tail=FALSE)

aic.white<-numeric()
for(i in 1:length(trees)) aic.white[i]<-whiteFit[[i]]$Trait1$aic


#-----------------------------------------------------
#   COMPARE ALL MODELS
#-----------------------------------------------------

# One way: use likelihood ratio test to compare all models to Brownian model

d.all<-cbind(d.lambda, d.delta, d.kappa, d.ou, d.eb, d.white)
p.all<-cbind(p.lambda, p.delta, p.kappa, p.ou, p.eb, p.white)

cat("Trait\tlambda\tdelta\tkappa\tou\teb\twhite\n")

for(i in 1:length(trees)) {
	cat("Tr", i, "\t");
	for(j in 1:6) {
		cat(round(d.all[i,j],2));
		if(p.all[i,j]<0.05) cat("*");
		if(p.all[i,j]<0.01) cat("*");
		if(p.all[i,j]<0.001) cat("*");
		cat("\t");
	}
	cat("\n");
}

# Another way: use AIC

aic.all<-cbind(aic.brown, aic.lambda, aic.delta, aic.kappa, aic.ou, aic.eb, aic.white)
foo<-function(x) x-x[which(x==min(x))]
daic<-t(apply(aic.all, 1, foo))

rownames(daic)<-colnames(spurs)
colnames(daic)<-c("Brownian", "Lambda", "Delta", "Kappa", "OU", "EB", "White")

cat("Table of delta-aic values; zero - best model\n")
print(daic, digits=2)

# Tree transformation. This depends on which model provides the best fit for your data. In this
example, I will assume this is Pagel's Lambda. Now let's transform the trees using the lambda-
parameter, which needs to be extracted first:
lambda <- list()
for (i in 1:length(trees))  {
lambda[[i]] <- lambdaFit[[i]]$Trait1$lambda  }

# Now transform the trees:
```

```
lambdaTrees <- list()
for (i in 1:length(trees))  {
lambdaTrees[[i]] <- lambdaTree(trees[[i]], lambda[[i]])  }

# Save the trees in a file and reload them.
write.nexus(lambdaTrees, file = "treeBlockLambda.nex")
lambdaTrees <- read.nexus("treeBlockLambda.nex")
# A new analysis can continue from here!

# To test if the transformed trees really fit the data best, run 'trees <- lambdaTrees' and then perform the whole analysis again. Results (results will differ for different data):
#     Brownian Lambda Delta Kappa  OU EB White
#[1,]        0      2   1.8   2.0 2.0  2    36
#[2,]        0      2   1.7   2.0 2.0  2    38
#[3,]        0      2   1.7   1.9 2.0  2    37
#[4,]        0      2   1.5   2.0 1.9  2    40
# Conclusion: now the Brownian motion fits the data best! This indicates that lambda-transforming the trees is the best way to continue

# Test if the lambda-value calculated by fitContinuous is significantly different from a lambda of 1 or 0:
lambdaTrees0 <- list()
lambdaTrees1 <- list()
for (i in 1:length(trees))  {
lambdaTrees0[[i]] <- lambdaTree(trees[[i]], 0)
lambdaTrees1[[i]] <- lambdaTree(trees[[i]], 1)  }

brownFitx <- list()
brownFit1 <- list()
brownFit0 <- list()
for (i in 1:length(trees))  {
brownFitx[[i]] <-  fitContinuous(lambdaTrees[[i]], spurs, bounds=list(beta=c(0,10000)))
brownFit1[[i]] <-  fitContinuous(lambdaTrees1[[i]], spurs, bounds=list(beta=c(0,10000)))
brownFit0[[i]] <-  fitContinuous(lambdaTrees0[[i]], spurs, bounds=list(beta=c(0,10000)))  }

# Extract AIC values for new models
aic.brownx<-numeric()
for(i in 1:length(trees)) aic.brownx[i]<-brownFitx[[i]]$Trait1$aic

aic.brown0<-numeric()
for(i in 1:length(trees)) aic.brown0[i]<-brownFit0[[i]]$Trait1$aic

aic.brown1<-numeric()
for(i in 1:length(trees)) aic.brown1[i]<-brownFit1[[i]]$Trait1$aic

# Compare AIC values
aic.all<-cbind(aic.brownx, aic.brown0, aic.brown1)
foo<-function(x) x-x[which(x==min(x))]
daic<-t(apply(aic.all, 1, foo))
```

```
rownames(daic)<-colnames(spurs)
colnames(daic)<-c("Lambdax", "Lambda0", "Lambda1")

cat("Table of delta-aic values; zero - best model\n")
print(daic, digits=2)

# Results: the lambda calculated by fitContinuous fits the data better than lambda=0 or lambda=1

# Compare p-values:
d.lambda1<-numeric()
for(i in 1:length(trees)) d.lambda1[i]=2*(brownFitx[[i]]$Trait1$lnl-brownFit1[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.lambda1=pchisq(d.lambda1, 1, lower.tail=FALSE)

d.lambda0<-numeric()
for(i in 1:length(trees)) d.lambda0[i]=2*(brownFitx[[i]]$Trait1$lnl-brownFit0[[i]]$Trait1$lnl)

# Calculate p values assuming chi-squared distribution with 1 d.f.
p.lambda0=pchisq(d.lambda0, 1, lower.tail=FALSE)

# Compare lambdas using likelihood ratio test:
d.all<-cbind(d.lambda0, d.lambda1)
p.all<-cbind(p.lambda0, p.lambda1)

cat("Trait\tlambda0\tlambda1\n")

for(i in 1:length(trees)) {
	cat("Tr", i, "\t");
	for(j in 1:2) {
		cat(round(d.all[i,j],2));
		if(p.all[i,j]<0.05) cat("*");
		if(p.all[i,j]<0.01) cat("*");
		if(p.all[i,j]<0.001) cat("*");
		cat("\t");
	}
	cat("\n");
}
# Result: both lambda=0 and lambda=1 are significantly different than lambda=x, which indicates
we should use our calculated lambda value
```

MEDUSA

```
# Run MEDUSA with a species richness table (so you leave only one species per clade):

# Remove all species except one per clade:
taxa <- list()
except <- list()

for (i in 1:length(trees))  {
```

```
taxa[[i]] <- trees[[i]]$tip.label
except[[i]] <- c(match("aciculatum2282", taxa[[i]]), match("paniculatumAM80", taxa[[i]]),
match("alpinum3574",taxa[[i]]), match("nanum2b", taxa[[i]]),
 match("reniforme141", taxa[[i]]), match("frutetorum211", taxa[[i]]), match("longicaule206",
taxa[[i]]))
# This gives the position of the taxa you'd like to keep in the taxa. Separate these taxa from the
others:
taxa[[i]] <- taxa[[i]][-except[[i]]]
# Now remove all remaining taxa:
trees[[i]] <- drop.tip(trees[[i]], taxa[[i]])  }
# "Phy' is now a tree with 7 tips and 6 internal nodes!
# NB: removing a species twice will result in a removal of all species from the tree!

# Now for the MEDUSA analysis:

# Load a table with the amount of species per clade
richness <- read.table("TabelMEDUSA.csv", header = F)

MEDUSAout <- list()
medusaSum <- list()
for (i in 1:length(trees))  {
# MEDUSA:
MEDUSAout[[i]] <- MEDUSA(trees[[i]], richness)
# Analyse results:
medusaSum[[i]] <- medusaSummary(MEDUSAout[[i]], plotTree=FALSE)

      # Run MEDUSA with all species in the tree (corrected for double species and non-random
sampling):

# Run MEDUSA
MEDUSAoutFull <- list()
medusaSumFull <- list()

for (i in 1:length(trees))  {
MEDUSAoutFull[[i]] <- MEDUSA(trees[[i]])
medusaSumFull[[i]] <- medusaSummary(MEDUSAoutFull[[i]], plotTree=FALSE)  }

#                    Analysis of results:

N.models <- list()
y <- list()
for (i in 1:length(trees))  {
N.models[[i]] <- MEDUSAout[[i]]$modelSummary$N.Models
y[[i]] <- which(N.models[[i]][2]==2, arr.ind=TRUE)  }
# 'y' will give a list of trees without rate shift, which will need to be excluded from further analysis

except2 <- c(6, 10, 25, 63, 78, 98, 106, 117, 119, 147, 150, 186, 195, 200)
MEDUSAout <- MEDUSAout[-except2]

Rbackground <- list()
```

```
Rforeground <- list()

for (x in 1:length(MEDUSAout))  {
Rbackground[[x]] <- MEDUSAout[[x]]$models[[2]]$par[1,1]
Rforeground[[x]] <- MEDUSAout[[x]]$models[[2]]$par[2,1]  }

fBackground <- unlist(Rbackground)
fForeground <- unlist(Rforeground)
x <- fForeground-fBackground
hist(x, breaks=30, main ="Difference in birth rate between different parts of the tree", xlab="Birth-rate after rate shift minus birth-rate before rate shit")

# Node where split occurs:
node <- list()
for (i in 1:length(MEDUSAout))  {
node[[i]] <- MEDUSAout[[i]]$models[[2]]$split.at[[2]]  }
```

QuaSSE

```
p <- starting.point.quasse(phy, spurs)

# linear function: linear in range xr[1]-xr[2], flat outside this range
# It probably makes most sense to set the boundaries of xr to 0 (shortest spur) and 10:
xr <- c(0, 10)
linear.x <- make.linear.x(xr[1], xr[2])

## NB: the 'sampling.f' scalar allows you to enter the estimated proportion of included species in the phylogeny. Important!
make.Pelargonium <- function(lambda, mu) make.quasse(phy, spurs, spurs.sd, lambda, mu)

# Set drift to 0
nodrift <- function(f) constrain(f, drift ~ 0)

# likelihood functions with varying functions for speciation (extinction remains constant)
f.speciation.constant <- make.Pelargonium(constant.x, constant.x)
f.speciation.linear <- make.Pelargonium(linear.x, constant.x)
f.speciation.sigmoid <- make.Pelargonium(sigmoid.x, constant.x)
f.speciation.humpshaped <- make.Pelargonium(noroptimal.x, constant.x)

# ML analysis, fitting the constant model first
control <- list(parscale = 0.1, reltol = 0.001)
mle.speciation.constant <- find.mle(nodrift(f.speciation.constant), p, lower = 0, control = control, verbose = 0)

# starting points for the constrained analyses based on this constrained fit
p.speciation.constant <- mle.speciation.constant$par
p.speciation.linear <- c(p.speciation.constant[1], l.m = 0, p.speciation.constant[2:3])
p.speciation.sigmoid <- p.speciation.humpshaped <- c(p.speciation.constant[1],
p.speciation.constant[1], mean(xr), 1, p.speciation.constant[2:3])
names(p.speciation.sigmoid) <- argnames(nodrift(f.speciation.sigmoid))
```

```
names(p.speciation.humpshaped) <- argnames(nodrift(f.speciation.humpshaped))

# ML analyses for the other functions (linear, sigmoid, and humpshaped)
mle.speciation.linear <- find.mle(nodrift(f.speciation.linear), p.speciation.linear, control = control,
verbose = 0)
mle.speciation.sigmoid <- find.mle(nodrift(f.speciation.sigmoid), p.speciation.sigmoid, control =
control, verbose = 0)
mle.speciation.humpshaped <- find.mle(nodrift(f.speciation.humpshaped),
p.speciation.humpshaped, control = control, verbose = 0)

# compare the fits of the different function. The constant speciation function ("full") is the default
function (all others are compared against this fit).
anova(mle.speciation.constant, linear = mle.speciation.linear, sigmoid = mle.speciation.sigmoid,
humpshaped = mle.speciation.humpshaped)

# run the fits with the drift parameter added, starting from the constrained model's ML parameters.
mle.drift.speciation.linear <- find.mle(f.speciation.linear, coef(mle.speciation.linear, TRUE), control
= control, verbose = 0)
mle.drift.speciation.sigmoid <- find.mle(f.speciation.sigmoid, coef(mle.speciation.sigmoid, TRUE),
control = control, verbose= 0)
mle.drift.speciation.humpshaped <- find.mle(f.speciation.humpshaped,
coef(mle.speciation.humpshaped, TRUE), control = control, verbose = 0)

# compare all the models.
anova(mle.speciation.constant, linear = mle.speciation.linear, sigmoid = mle.speciation.sigmoid,
humpshaped = mle.speciation.humpshaped,
drift.linear = mle.drift.speciation.linear, drift.sigmoid = mle.drift.speciation.sigmoid,
drift.humpshaped = mle.drift.speciation.humpshaped)

# Drift parameter of the different model fits. A positive parameter means an increase in spur length.
c(linear = coef(mle.drift.speciation.linear)[["drift"]], sigmoid = coef(mle.drift.speciation.sigmoid)
[["drift"]],
humpshaped = coef(mle.drift.speciation.humpshaped)[["drift"]])

# Add node names:
phy$node.label <- paste("nd", 1:phy$Nnode, sep="")

# Make QuaSSE objects with a split:
f.speciation.constant.constant <- make.quasse.split(phy, spurs, spurs.sd, constant.x, constant.x,
"nd3", Inf, sampling.f = 0.72)
# NB: The right node still has to be specified!

#  Constrain drift to be zero and assume that both partitions have the same diffusion coefficient.
g.speciation.constant.constant <- constrain(f.speciation.constant.constant, drift.1 ~ 0, drift.2 ~ 0,
diffusion.2 ~ diffusion.1)

# Generate a starting point from the single partition ML point:
p.speciation.constant.constant <- c(p.speciation.constant, p.speciation.constant[1:2])
names(p.speciation.constant.constant) <- argnames(g.speciation.constant.constant)
```

```
# Now the split function should have basically the same likelihood as the single partition function:
mle.speciation.constant$lnLik - g.speciation.constant.constant(p.speciation.constant.constant)
# (The result of this subtraction should be close to zero.)

# Run the ML search:
mle.speciation.constant.constant <- find.mle(g.speciation.constant.constant,
p.speciation.constant.constant, control = control, lower = 0, verbose = 0)

# Now repeat this for linear speciation functions:
f.speciation.linear.linear <- make.quasse.split(phy, spurs, spurs.sd, linear.x, constant.x, "nd3", Inf,
sampling.f = 0.72)
g.speciation.linear.linear <- constrain(f.speciation.linear.linear, drift.1 ~ 0, drift.2 ~ 0, diffusion.2 ~
diffusion.1)
g.speciation.linear.constant <- constrain(g.speciation.linear.linear, l.m.2 ~ 0)
g.speciation.constant.linear <- constrain(g.speciation.linear.linear, l.m.1 ~ 0)

# Genearate starting points:
p.speciation.constant.constant <- coef(mle.speciation.constant.constant)
p.speciation.linear.linear <- c(p.speciation.constant.constant[1], 0,
p.speciation.constant.constant[2:4], 0, p.speciation.constant.constant[5])
names(p.speciation.linear.linear) <- argnames(g.speciation.linear.linear)

# Run ML search:
mle.speciation.linear.linear <- find.mle(g.speciation.linear.linear, p.speciation.linear.linear, control =
control, verbose = 0)

# Starting points for the partial functions:
p.speciation.linear.constant <- c(coef(mle.speciation.linear.linear)[1:3],
p.speciation.linear.linear[c(4,5,7)])
p.speciation.constant.linear <- c(p.speciation.linear.linear[c(1,3,4)],
coef(mle.speciation.linear.linear)[5:7])

# Run the ML searches:
mle.speciation.linear.constant <- find.mle(g.speciation.linear.constant, p.speciation.linear.constant,
control = control, verbose = 0)
mle.speciation.constant.linear <- find.mle(g.speciation.constant.linear, p.speciation.constant.linear,
control = control, verbose = 0)

# Now compare all the models:
anova(mle.speciation.constant, linear = mle.speciation.linear, sigmoidal = mle.speciation.sigmoid,
humpshaped = mle.speciation.humpshaped, part.constant =
mle.speciation.constant.constant, part.linear.background = mle.speciation.linear.constant,
part.linear.foreground = mle.speciation.constant.linear, part.linear =
 mle.speciation.linear.linear)

# Model parameters can be assessed by entering the name of the model
```

Appendix D - Extra activities

Presentations

- Project presentation, Trends from Trees symposium, Wageningen University, October 6, 2011.
- Project presentation, the Department of Botany, University of Cape Town, November 16, 2011.
- Presentation about morphological analyses in R and other programs, journal club, Biosystematics. Group, Wageningen University, March 16, 2012.
- Presentation about morphological analyses in R, Current Trends in Phylogenetics PhD-course, Wageningen University, October 25, 2012.

Collections

- For two species (*P. longicaule* and *P. rapaceum*) I collected leaf samples of 50 individuals, belonging to 10 populations (evenly divided per species), for future DNA analysis.

Miscellaneous

- I provided information about floral characteristics and putative pollinators of South African *Pelargonium* species for a project with Jonathan Colville, John Manning, and Timo van der Niet.