Molecular Phylogenetics and Evolution 137 (2019) 33–43

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Plastome based phylogenetics and younger crown node age in *Pelargonium*

Sara J. van de Kerke[a,*], Bikash Shrestha[b], Tracey A. Ruhlman[b], Mao-Lun Weng[b,d], Robert K. Jansen[b,e], Cynthia S. Jones[c], Carl D. Schlichting[c], Samin Hosseini[a,f], Setareh Mohammadin[a], M. Eric Schranz[a], Freek T. Bakker[a,*]

[a] Biosystematics, Plant Science Group, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, the Netherlands
[b] Department of Integrative Biology, University of Texas at Austin, 205 W. 24th St. Stop C0930, TX 78712, USA
[c] Department of Ecology & Evolutionary Biology, University of Connecticut, 75 N. Eagleville Road, Storrs, CT 06269, USA
[d] Department of Biology, Westfield State University, Westfield, MA 01086, USA
[e] Department of Biological Sciences, Faculty of Science, King Abdulaziz University, P.O. Box 80141, Jeddah 21589, Saudi Arabia
[f] Department of Plant Protection, College of Agriculture, Vali- e-Asr University of Rafsanjan, Rafsanjan, Iran

ARTICLE INFO

Keywords:
Pelargonium
Geraniaceae
Phylogeny
Plastome
Time-calibrated

ABSTRACT

The predominantly South-African plant genus *Pelargonium* L'Hér. (Geraniaceae) displays remarkable morphological diversity, several basic chromosome numbers as well as high levels of organelle genomic rearrangements, and represents the 7th largest Cape Floristic Region clade. In this study, we reconstructed a phylogenetic tree based on 74 plastome exons and nuclear rDNA ITS regions for 120 species, which represents 43% taxon coverage for *Pelargonium*. We also performed a dating analysis to examine the timing of the major radiations in the genus.

Phylogenetic analyses of nucleotide, amino acid, and ITS alignments confirmed the previously-documented subgeneric split into five main clades ((C1,C2),(B(A1,A2))) although clade only A1 received low bootstrap support.

Using calibration evidence from a range of sources the *Pelargonium* crown age was estimated to be 9.7 My old, much younger than previous estimates for the genus but similar to recent studies of other Cape Floristic lineages that are part of both Fynbos and Succulent Karoo biomes.

1. Introduction

The predominantly South African genus *Pelargonium* L'Hér. (Geraniaceae) is morphologically diverse both in life forms, ranging from herbaceous annuals, woody (sub)shrubs, geophytes, rosette herbs to stem succulents and by remarkable variation in floral and leaf morphology (Bakker et al., 2005, 1999; Jones et al., 2009, 2003; Nicotra et al., 2008; Röschenbleck et al., 2014). *Pelargonium* is also characterised by extensive genomic variability, with six different basic chromosome numbers (Bakker et al., 2005), substantial variation in nuclear genome size (Weng et al., 2012), the independent occurrence of several polyploid series and unprecedented levels of variation in organelle genomes (Blazier et al., 2011, 2016b; Park et al., 2015; Bakker et al., 2006; Chumley et al., 2006; Guisinger et al., 2008, 2011; Mower et al., 2007; Parkinson et al., 2005; Weng et al., 2012, 2013). In addition, many species exhibit biparental inheritance and cytonuclear incompatibility (Ruhlman and Jansen, 2018). *Pelargonium* is one of a handful of speciose Cape lineages that span multiple biomes in Greater Cape Floristic Region, making this clade a promising model system for testing ecological and evolutionary hypotheses (e.g. Moore et al., 2018; Verboom et al., 2009).

Of the ~280 species of *Pelargonium*, approximately 200 species occur in the Greater Cape Floristic Region (GCFR) in South Africa (Linder, 2003; Manning and Goldblatt, 2012; Snijman, 2013) and have been well-documented taxonomically (Van der Walt and Vorster, 1981, 1988). Morphological, palynological, phytochemical and karyological data have been used in an extensive range of taxonomic studies (e.g. Van der Walt et al., 1990; Van der Walt and Vorster, 1983, 1981; see Röschenbleck et al., 2014 for an overview). New species are still being described, especially in the geophytic sect. *Hoarea* (Manning et al., 2015; Marais, 2016), which was considered a non-adaptive radiation nested within an adaptive radiation by Bakker et al. (2005). Their appeal as garden plants (dating back to Victorian times, Sweet (1822)) has led to interspecific crosses of *Pelargonium* that have resulted in a wide variety of highly valued commercial cultivars (Albers & van der Walt, 2007), especially in the *P.* × *hortorum* hybrid complex, *P. peltatum*, *P.*

Abbreviations: POC clade, *Polyactium-Otidia-Cortusina* clade; GCFR, Greater Cape Floristic Region

* Corresponding authors.
E-mail addresses: sara.vandekerke@wur.nl (S.J. van de Kerke), freek.bakker@wur.nl (F.T. Bakker).

https://doi.org/10.1016/j.ympev.2019.03.021
Received 10 July 2018; Received in revised form 23 March 2019; Accepted 25 March 2019
Available online 26 March 2019
1055-7903/

cucullatum, and *P. tricolor*, and some species such as *P. citronellum* and *P. graveolens* are important for essential oil production (Blerot et al., 2015).

Phylogenetics of *Pelargonium* has been investigated in a series of studies spanning the last three decades. Price and Palmer (1993) published the first DNA-based generic-level phylogenetic tree of Geraniaceae and found *Pelargonium* to be sister to the rest of the family (except *Hypseocharis*) and inferred a subgeneric split that correlated with chromosome size. Subsequent phylogenetic studies using increased taxon sampling confirmed this pattern for both the internal transcribed spacer (ITS) of nuclear ribosomal DNA (rDNA) and the plastid *trnL-F* group I intron and *trnL-F* spacer regions for small and large chromosome species (Bakker et al., 1999, 1998; Jones and Price, 1995). Making use of the remarkably elevated substitution rates in Geraniaceae mitochondrial DNA (mtDNA), Bakker et al. (2000b) found species-level phylogenetic resolution using exons 4 and 5 of mitochondrial-encoded *nad*1, which is otherwise fairly conserved across angiosperms. Using markers from all three genome compartments (*nad*1, *trnL-F*, and rDNA ITS) for 149 *Pelargonium* species (i.e. 53% taxonomic sampling), Bakker et al. (2005, 2004b) inferred a phylogeny and proposed a (C(B,A)) division of the genus with A and C clades each divided into two main clades. This analysis used weighted parsimony and heavily relied on non-coding sequences and recoded insertions and deletions (indels) (Bakker et al., 2004), which were found to provide ~20% of the total plastid DNA signal (Bakker et al., 1999). The ((C1,C2),(B(A1,A2))) *Pelargonium* phylogenetic pattern was later confirmed using Bayesian Inference on the same data (Jones et al., 2009).

Weng et al. (2012) noted the lack of overlap in taxonomic sampling of the different gene sequences in previous studies and recommended expanded gene and taxon sampling. The authors partially fulfilled this need by adding additional gene sequence data for the markers nad5, *ndhF*, *rbcL*, *mat*K, *rpo*C1, in addition to *trnL-F*. However, *Pelargonium* taxon sampling remained limited (only 21%) in that study. By adding the atp*B-rbcL* spacer to an increased taxon sampling for the *trnL-F* data set of Bakker et al. (2004), Röschenbleck et al. (2014) achieved a taxonomic coverage of 38% and confirmed a ((C1,C2),(B(A1,A2))) (but not A1 and A2) topology. However, nuclear and mitochondrial data were not included. Röschenbleck et al. (2014) proposed raising these four clades to the subgenus level. These subgenera were then further divided into sixteen sections, most of which confirmed the previously-proposed sectional classification (Bakker et al., 2004), and are supported by floral morphological data (Röschenbleck et al., 2014), although a general morphology based key has not been developed for *Pelargonium*.

Phylogenetic studies focusing on single *Pelargonium* clades have been performed for Clade C (James et al., 2004, based on plastome RFLPs), sect. *Hoarea* (Touloumenidou et al., 2003, based on rDNA ITS sequences), clade B (Bakker et al., 1998) and the Australian clade (Nicotra et al., 2016, using a population genomic approach). For sect. *Otidia,* an AFLP approach was used to study relationships in the *P. carnosum – P. paniculatum* and *P. alternans* complexes (Becker and Albers, 2010, 2009).

Four molecular dating studies have been performed in *Pelargonium,* but the results are contradictory. In 2005, using r8s and the non-parametric rate smoothing approach on *trnL-F* and nuclear rDNA ITS sequences, Bakker et al. estimated the *Pelargonium* crown node to have originated around 30 Mya. Based on the same data, but using BEAST analyses, Verboom et al. (2009) came to a similar estimate of 34.54 My, making it the oldest Fynbos biome clade giving rise to Succulent Karoo clades (Verboom et al., 2009). The deepest split into clades A,B versus C (with different chromosome size), would coincide with climatic changes in the Oligocene/Miocene (Goldblatt et al., 2002). Additionaly, the estimated age of the winter-rainfall clade A2 of 22 My coincides with the Early Miocene and was interpreted to be linked with the emergence of summer drought by Bakker et al. (2005). During the Miocene-Pliocene climate change around 10 Mya, upwelling of the cold Benguela current was established and, in combination with a strengthened South Atlantic pressure cell, could well have caused dry summers along the west coast of southern Africa. This is considered to be associated with the formation of the Succulent Karoo biome and the radiation of new, xerophytic, *Pelargonium* species (Bakker et al., 2005; Linder, 2003; Verboom et al., 2009).

In contrast, both Fiz et al. (2008) and Palazzesi et al. (2012) estimated the date for the *Pelargonium* crown node to be 10–15 Mya, corresponding with a transition to a drier and colder climate in the mid-Miocene (Linder, 2003). These studies both used fossilized pollen as evidence for their analyses (see Supplementary Table S1 for an overview of these studies). Given the discrepancies among *Pelargonium* dating studies, an update using increased numbers of characters and taxa is needed.

The aims of this study are to resolve phylogenetic relationships within *Pelargonium* using extended character sampling (74 plastome protein-coding genes as well as nuclear rDNA ITS) and to provide robust age estimate. Previous *Pelargonium* phylogenetic studies have relied on plastome intergenic and Group I intron sequences (such as *trnL-F* and the *atpB-rbcL* spacer), in addition to indels in the same spacers. Our approach using exons provides sufficient data to resolve remaining phylogenetic issues in *Pelargonium*, and results in an improved phylogenetic framework for future genomic, morphological and evolutionary studies.

2. Material and methods

2.1. Taxon sampling

Leaf material was obtained from various sources (see Supplementary Table S2), including plants obtained in the field and from botanical gardens. One collection series was sampled from various wild populations collected across South Africa and was silica gel dried by Schlichting, Jones and collaborators during 2012–2015. Vouchers are deposited at CONN. Another series comes from the Jansen lab with material obtained from Geraniaceae.com. Plants were maintained in the University of Texas at Austin greenhouse and vouchers were made for each species and deposited in TEX-LL. Based on previous phylogenetic studies, we selected samples to represent all five major clades within *Pelargonium*. This resulted in a combined set of 148 accessions representing 120 species (approximately 43% of all *Pelargonium* species). We chose two accessions of *Hypseocharis biloba* (NC_023260.1; Bakker et al., 2016) as outgroups.

2.2. DNA isolation, Illumina sequencing and plastome assembly

DNA from the wild collected samples was isolated from silica gel dried leaf material in the Bakker lab using a modified CTAB protocol (Bakker et al., 1998; Doyle, 1991) after grinding in liquid nitrogen. Following Isopropanol precipitation, the Wizard® DNA Clean-Up System (Promega, 2010) was used to further purify the samples. DNA quantity was determined using a Qbit spectrophotometer (Thermo Fisher Scientific, 2008). Samples yielding > 20 ng were shipped to BGI Hong Kong for library preparation and Illumina paired end (PE) sequencing. A few accessions with low total yield (between 20 and 50 ng) underwent a whole genome amplification step.

For the plastome dataset, Illumina PE reads were assembled using IOGA, an automated bioinformatics pipeline (Bakker et al., 2016), which uses both *de novo* and reference-based assembly, by mapping reads against a panel of reference genomes that need not be closely related to the target. As a reference library, the same reference plastomes as in Bakker et al. (2016) were used, including complete plastomes of *P. alternans* (NC_023261.1) and *P.* × *hortorum* (DQ897681.1). IOGA uses SOAPdeNovo (Xie et al., 2014) in order to assemble mapped reads into contigs. A range of k-mer sizes was used (33, 55, 75, 95) to optimise the assembly, assuming an insert size of 250 bp. Plastome-

derived reads remaining in the initial total read pool that overlapped with the assembled contigs were mapped to the contigs and assembled *de novo*. New iterations of mapping and assembly were then performed until no new reads could be added to the contigs. Final assembly, usually producing a range of contigs, was performed using SPADES (Bankevich et al., 2012) as implemented in IOGA, followed by selection of candidate assemblies using either Assembly Likelihood Estimation (ALE) score (Clark et al., 2013), overall coverage or N50. When plastome read coverage was exceptionally high for a particular sample (i.e. > 1000), a subsample of one to five million reads was taken before re-assembly.

The methods for DNA isolation, Illumina sequencing, assembly and annotation for the 61 species contributed from the Jansen collection are described in Blazier et al. (2016a) and Weng et al. (2014). For 21 of the 61 *Pelargonium* species, complete plastomes were completed and 74 protein coding genes were extracted. For the remaining species, the genes were extracted from contigs of draft genome assemblies.

For the nuclear rDNA dataset, both the Bakker and Jansen collection underwent the IOGA assembly procedure as described above. A collection containing all previously published rDNA ITS accessions for *Pelargonium* available in GenBank was used as reference.

2.3. Annotation and gene selection

All plastid assemblies from the Bakker lab were annotated in Geneious 8.1.6 (Kearse et al., 2012) using *P. alternans* (Weng et al., 2013) as reference and setting the sequence similarity threshold at 75%. In total, a set of 74 protein coding genes was extracted from the assembly data and gene alignments were compiled. Each gene alignment was split into separate intron/exon alignments with the use of the TAIR webtool (https://www.arabidopsis.org/index.jsp; Supplementary Table S3 and S4). The complete plastome sequence of *Hypseocharis biloba* (NC_023260.1) chosen as outgroup underwent the same procedure of gene extraction and alignment.

All nuclear rDNA assemblies were annotated using *Brassica rapa* (KM538956.1) as reference for the ribosomal (18S, 5.8S, and 26S) as well as ITS1 and ITS2. In cases when not all components of the rDNA region could be retrieved in one piece, we used the universal primers as designed by White et al. (1990) to find the ITS1 and ITS2 boundaries.

2.4. Alignment and data matrix construction

For both plastome and nuclear data, MAFFT v. 7 was used for optimising each alignment under 'auto' settings (Katoh and Standley, 2013). All alignments were visually inspected in Mesquite v. 3.04 (Maddison and Maddison, 2015) and manually adjusted where needed. Reading frames were set for all coding region alignments using the 'Minimize stop codons' function. Alignments were trimmed accordingly to reading frame in order to eliminate incomplete codons. A 'Plastome Introns and Exons' (PIE) matrix included all above described alignments, concatenated using SequenceMatrix (Vaidya et al., 2011) into a single alignment. In addition, a 'Plastome Exons AminoAcid' (PE-A) matrix contained an amino-acid version of the exon-only data. The number of parsimony informative sites was calculated using PAUP* (Swofford, 2002).

2.5. Phylogenetic analyses

Maximum likelihood-based phylogenetic analysis of plastome matrices was performed using RAxML v. 8.2.8 on the XSEDE supercomputer at the CIPRES Science Gateway platform (Miller et al., 2010; Stamatakis, 2014). Two partition schemes for the PIE matrix were compared: 1. unpartitioned and 2. partition assigned by PartitionFinder v. 1.1.1 (Lanfear et al., 2012), which selects from alternative gene- or codon-position level partitioning on the basis of the Bayesian Information Criterion (BIC). The PE-A matrix was analysed under an unpartitioned model (using the PROTGAMMADAYHOFF amino acid substitution model) as optimising AA models in multiple partitions is computationally prohibitive. RAxML analyses of DNA sequence data was performed using the GTR + GAMMA model. All analyses included inference of the 'best tree' as well as generation of 1000 bootstrap trees, to obtain node support measures. In addition, we used MrBayes v. 3.2.6 (Huelsenbeck and Ronquist, 2001; Ronquist et al., 2012) for a Bayesian inference of our plastome alignments (500 million generations, nruns = 2, four chains, sampled every 30.000th generation, nst = mixed, temp = 0.05/0.2).

Phylogenetic analysis of the rDNA ITS matrix was performed under ML using IQ-TREE with standard settings on the IQ-TREE web server (iqtree.cibiv.univie.ac.at) generating 1000 bootstrap trees (Hoang et al., 2018; Kalyaanamoorthy et al., 2017; Nguyen et al., 2015; Trifinopoulos et al., 2016). The analysis includes Ultrafast model selection (ModelFinder) and Ultrafast bootstrap (UFBoot).

All resulting phylogenetic trees were visualised using TreeGraph2 (Stöver and Müller, 2010).

2.6. Divergence date estimates

We used BEAST v1.8.4 (Drummond et al., 2012) to infer a time-calibrated phylogenetic tree of *Pelargonium* using the PIE matrix, adding fourteen Geraniales genomes in order to accommodate all available fossil calibrations (Supplementary Table S5). We used three calibration methods: (1) Fossil calibration, using estimated ages of available fossils of *Geranium*, *Erodium*, Vivianaceae and *Pelargonium* set with a log-normal distribution for each calibration prior which has an unbound tail reflecting the uncertainty of the maximum age of the node (Ho and Phillips, 2009; Palazzesi et al., 2012). (2) Secondary calibration, in which the crown node age corresponding to Geraniales as estimated by Wang et al. (2009) was used to calibrate our phylogenetic tree using a normal prior. (3) Ecological calibration, in which we assumed that clade A2 (the 'Winter Rainfall Region' clade in Bakker et al., 2005) could have emerged in response to the establishment of the Mediterranean type climate in the South Western Cape, which has been estimated as late-Miocene: the tertiary fossil record of southern Africa suggests that the earliest summer-drought conditions became established approximately 10 Mya (Linder, 2003). We therefore calibrated the A2 node with this age using a normal prior distribution. In order to assess possible calibration incongruence, we explored the following calibration combinations: Fossils, 2nd, and Ecological calibration separate, the combination of Fossils and 2nd calibration, the combination of 2nd and Ecological calibration, and all three methods combined (Table 1).

We used the uncorrelated lognormal relaxed molecular clock (UCLD) models to account for rate variability among lineages and chose the Yule speciation model, which is considered the most appropriate model for species-level datasets (Bouckaert et al., 2014). We set the prior distribution for mean rate of the clock model as recommended by Ferreira and Suchard (2008) and used the GTR substitution model and assumed site rates to be 4Γ distributed (as suggested by IQ-tree, not shown). We performed one MCMC analysis per dating scenario of 400 million generations each, sampling every 10.000 steps. For the scenario combining all calibration methods, we performed four additional MCMC analyses.

We combined log and tree files using LogCombiner v.1.8.4. (Drummond et al., 2012) and checked for convergence using VMCMC (Visual Markov chain Monte Carlo, Ali et al., 2017) to diagnose global convergence of the whole MCMC chain to the target distribution by calculating Gelmen-Rubin and Gewek parameters (Ali et al., 2017). Split frequency plots that measure topological differences among chains were generated in RWTY (Warren et al., 2017). In case of appropriate convergence, frequencies in the cumulative plot should level off, indicating that clade/split is present in both posterior distributions. We used TreeAnnotator v1.8.4 (implemented in BEAST tools package) with

Table 1
Prior setting for calibration evidence for different calibration combinations.

Node	Fossils		2nd		Ecological	
	Age[1]	Prior	Age	Prior	Age	Prior
GAL			99-109	Normal 104 (2)[3]		
GER	7.25 (0.005)	Lognormal 7.25 (4, 7.24)[2]				
E	7.25 (0.005)	Lognormal 7.25 (4, 7.24)[2]				
MFV	10 (0.3)	Lognormal 10 (2.5, 9.7)[2]				
GexH	28.4 (0.1)	Lognormal 28.4 (3,28.3)[2]				
A2					8-10	Normal 9 (0.4)[3]
	Combination I		Combination II		Combination III	
	Combination IV					
			Combination V			
	Combination VI					

[1] Age range in Mya.
[2] Mean, sd, hard minimum bound (= *offset)* (in real space).
[3] Mean, sd; GER (*Geranium*), E (*Erodium*), MFV (*Melianthus-Franoa-Viviania*), GexH (Geraniales excluding *Hypseocharis*), A2 (*Pelargonium* clade A2) and GAL (Geraniales) indicate calibrating node.

a burn-in of 10% to summarize the tree results.

3. Results

3.1. Assembly and alignments

In the Bakker lab, 80 new *Pelargonium* specimens were sampled for DNA extraction and Illumina sequencing in this study. After library preparation and sequencing, the total number of reads ranged from 5,286,525 (*P. minimum*) to 29,102,984 (*P. saxifragoides*). For the specimens from the Jansen lab sequencing depth was much higher, around 60 M reads each (see Supplementary Table S6). Average assembly size of the plastomes was 154,624 bp with an average read coverage of 690 after sub-sampling for the Bakker lab samples and on average over 1500× for those from the Jansen lab. The total concatenated PIE alignment (Plastome Introns and Exons) was 64,388 bp in length with 6305 (9.8%) potentially parsimony informative sites, and covered 43% of all known *Pelargonium* species, whereas the PE-A matrix contained 18,800 amino acid residues of which 3187 (17%) were parsimony informative. All sequences have been deposited in GenBank (Supplementary Table S7) and the final PIE and ITS alignments and resulting phylogenetic trees can be found in Supplementary File S8 and S9 and under TreeBase Submission ID 24185.

3.2. Phylogenetic patterns

PartitionFinder analysis suggested the data be partitioned over 22 different partitions (Supplementary Table S10 and S11) that corresponded to codon position rather than gene functional group as in Guisinger et al. (2008).

Comparisons of ML tree topologies and support values for the unpartitioned PIE matrix, the partitioned PIE matrix and the unpartitioned PE-A matrix detected few topological discrepancies (indicated by * in Fig. 1): five within clade C1, five within clade B, two within the sect. *Pelargonium* and one within the *Hoarea* clade. When bootstrap support values differed, higher values were generally obtained for the unpartitioned PIE data set. Tree topologies for MrBayes analyses were congruent with those for RAxML. The phylogenetic tree inferred from the nuclear rDNA ITS matrix produced the same topology as the plastome matrices for major clades of *Pelargonium* although the topologies of species within clades were different (Fig. 2). For example, based on plastome sequences *P. plurisectum*, together with *P. barklyi, P. articulatum*, and *P. alchemilloides*, is confidently placed as sister to the remainder of the clade corresponding to the section *Ciconium* while the nuclear rDNA patterns suggests *P. plurisectum* to be placed more central in the clade. The same small species-level shift within clades occurred for *P. cucculatum, P. cordifolium, P. capitatum, P. glutinosum, P. ionidiflorum, P. alchemilloides,* and *P. wuppertalense*. Larger incongruences occur for *P. klinghardtense* which shifts from the clade corresponding to its taxonomic section in the plastome based phylogeny to sister species of the section *Magnistipulacea* based on nuclear rDNA. The reverse is the case for *P. desertorum* whose position is likewise flexible but also has the status of 'unassigned species' within the subgenus *Pelargonium* (Röschenbleck et al., 2014). Also *P. panduriforme* ends up in the 'wrong' clade based on nuclear rDNA sequences. The trio *P. gibbosum, P. crithmifolium*, and *P. crassicaule* group together between the clades of their respective taxonomic sections, albeit based on rather low bootstrap support.

3.3. Divergence date estimates

In order to compare the calibration results for the single MCMC runs based on different combinations of calibration methods, we focus on the *Pelargonium* main crown nodes (Fig. 3), i.e. 'Winter-rainfall' clade A, clade A1, 'xerophytic' clade A2, clade B and clade C . The estimated age of *Pelargonium* nodes for the Fossil and Ecological calibration separately are overall comparable, while the 2nd calibration methods shows quite a different pattern with much older age estimates. In addition, the range of HPD's is much larger. When combining the Fossil and 2nd calibration methods, age estimates appear to be predominantly influenced by the 2nd calibration evidence. The combination of 2nd and Ecological calibration does not appear to be subject to this influence as results are comparable with the separate Ecological calibration results. The result of the combination of all three calibration methods fits in this pattern with mean age estimates again congruent with the Ecological calibration.

We consider the scenario including fossil, secondary, as well as ecological calibration as the final result because it is based on maximum evidence (Fig. 4). Based on the four extra MCMC runs for this scenario, Geraniaceae crown node appears to have proliferated in the Middle Eocene (~35.8, 95% HPD = 29.5–45.1 Mya) with the *Pelargonium* crown node proliferating in the Late Miocene (9.7 Mya, 95% HPD = 9.0–10.5 Mya). Based on our results, the crown of the oldest clade of *Pelargonium*, clade C, diverged around 8.6 Mya (95% HPD = 7.5–9.7 Mya) while the diversification of B, A1 and 'xerophytic' A2 occurred in the Early Pliocene and the Late Miocene (4.5 Mya, 95% HPD = 2.7–6.3 Mya, 4.5 Mya, 95% HPD = 2.8–6.2 Mya, and 5.3 Mya, 95% HPD = 3.9–6.7 Mya, respectively).

4. Discussion

Pelargonium has been the focus of an expanding series of

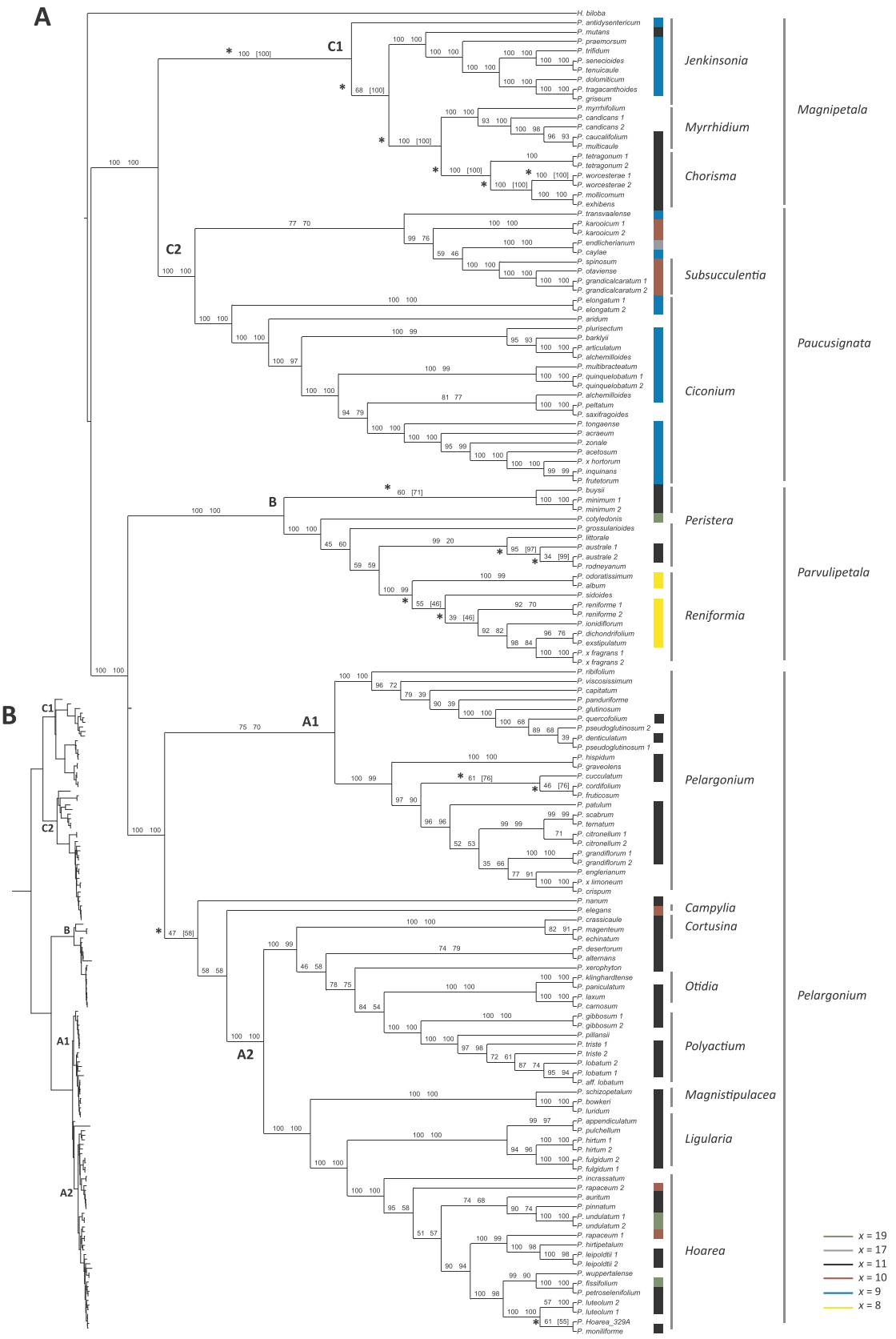

Fig. 1. (A) RAxML tree based on matrix PIE (unpartitioned) and PE-A matrices (GTR + GAMMA), in cladogram style. Bootstrap values indicate support at node for PIE/PE-A analysis respectively. Brackets indicate conflict between analyses. Clade labels sensu section- and subgenus level classification of Röschenbleck et al. (2014). Capital letters correspond to main clades. (B) Same tree as A showing branch lengths in nucleotide substitutions per site (outgroup pruned from tree).

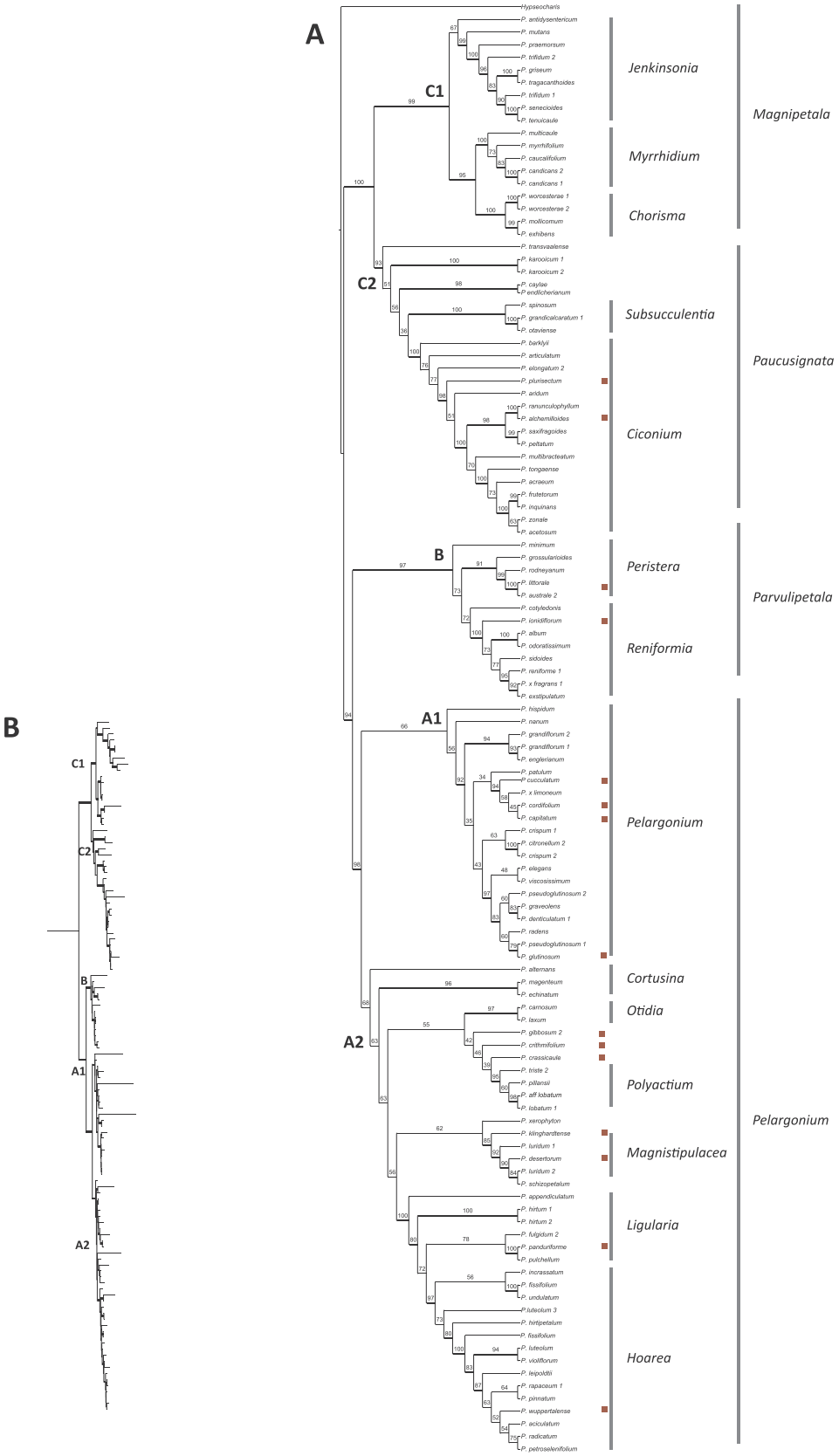

Fig. 2. (A) RAxML tree based on ITS matrix (IQ-TREE), with bootstrap values indicated. Clade labels sensu section- and subgenus level classification of Röschenbleck et al. (2014). Red squares indicate species-level plasto-ribo incongruence. Capital letters correspond to main clades. (B) Phylogram showing branch lengths in substitutions per site (outgroup pruned from tree) resulting from the RAxML analyses on ITS matrix (IQ-TREE). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

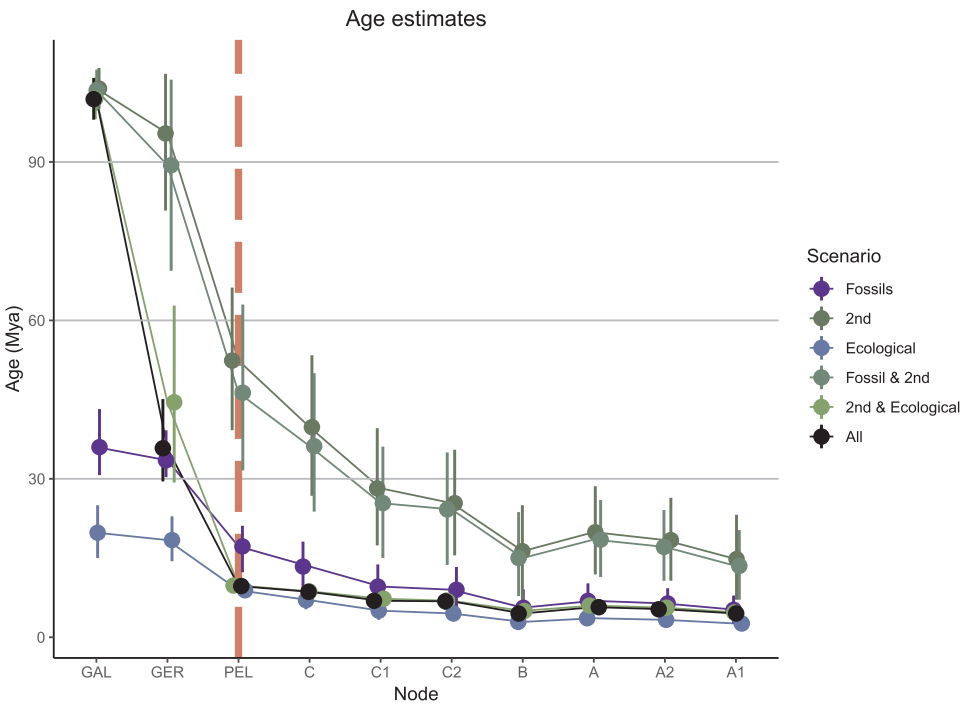

Fig. 3. Boxplot of divergence date estimations for main nodes, based on the six combinations of calibration evidence, coloured as indicated: Fossils, 2nd, and Ecological calibration separate, the combination of Fossils and 2nd calibration, the combination of 2nd and Ecological calibration, and all three methods combined. Error bars represent HPDs.

phylogenetic studies (Bakker et al., 2004, 1999, 1998; Jones et al., 2009; Price and Palmer, 1993; Röschenbleck et al., 2014; Weng et al., 2012). In those studies, an increasing number of phylogenetic markers has been utilized from all three genomic compartments and taxonomic coverage has been substantially expanded up to 53%. However, a common set of markers needed to link these studies has so far been missing, leaving unclear to what extent missing data in the phylogenetic matrices has been influential and whether inter-genomic topological incongruence may have occurred. We compiled a matrix of 74 plastid genes as well as the nuclear ITS region for 120 *Pelargonium* species, achieving 43% taxonomic coverage of the genus. Although our taxon coverage is far from complete, our extensive gene sampling includes species representing all previously reported main clades.

Bakker et al. (2004) found that incongruence between phylogenetic trees generated from nuclear rDNA and plastome sequences was limited to the species-level and occurred predominantly within clades that corresponded to previously described taxonomic sections. Our findings reveal the same pattern detected by Bakker et al. (2004): incongruence between phylogenetic trees generated from nuclear rDNA and plastome sequences is limited to the species-level and occurs predominantly within clades (Figs. 1 and 2). Overall, we feel there are no major incongruences between the plastome and nuclear rDNA perspective and, considering the relatively low bootstrap support values for the latter, decided focus on the plastome markers, leaving the nuclear and mitochondrial perspectives for future studies. Arguably, combining all genomic compartments in an overarching phylogenetic analysis would require a species tree estimation approach using multi-species coalescent methods (Liu et al., 2009).

We restricted our plastome-based analyses to predominantly protein-coding exon sequence data, ignoring fast-evolving spacer regions (i.e. the Plastome Exon and Intron (PIE) data partition). Spacer regions have been useful at the species level in *Pelargonium* and, for instance, the length variation present in the *trnL-F* regions (using indel coding) represented 20% of the phylogenetic signal (Bakker et al., 1999). However, the high frequency of rearrangements and indels observed in Geraniaceae plastomes (Guisinger et al., 2011) can confound homology assessment among sites in most spacer regions. Instead we relied here on the 6305 informative characters residing in the 74 plastome exons and 10 corresponding introns, and expect that adding additional spacer regions or indel characters would not significantly alter the results. The resulting tree topology for Bayesian and maximum likelihood analyses were congruent and overall comparable to previous studies (Bakker et al., 2004; Röschenbleck et al., 2014; Weng et al., 2012).

4.1. *Pelargonium* species – level patterns

Although in some phylogenetic studies deeper nodes appear to be well-supported, studies using larger taxon sampling (e.g. Bakker et al., 2004) show poor support for these nodes, suggesting the high-support for deep nodes in these low taxon sampling studies is artefactual. The subgeneric split into a small and large chromosome clade was confirmed here (Bakker et al., 2004; Price and Palmer, 1993; Van der Walt et al., 1990; Weng et al., 2012). We found 100% bootstrap support for four of the five major clades. Support was low for clade A1 (75/70%) similar to all previous studies, challenging its validity. Clade A1 includes species from sect. *Pelargonium* that are characterised by shrub and sub-shrub life-forms making them a well-defined clade morphologically. This clade is the type section for *Pelargonium* (De Candolle, 1824; Röschenbleck et al., 2014; Sweet, 1822; van der Walt, 1985) but apparently its distinctness is not supported by DNA data in all cases.

The position of *P. nanum* has been a longstanding issue in *Pelargonium* phylogenetics (Bakker et al., 1999). *Pelargonium nanum* has floral and vegetative morphology more typical of species in clade B, with small, bicolored flowers and an annual habit (Röschenbleck et al., 2014). However, over the years, it has been proposed as part of clade A1 (Bakker et al., 2004), sister to clade A2 (Weng et al., 2012) or sister to the entire clade A (Röschenbleck et al., 2014). The inclusion of *P. nanum* prevented Markov Chain convergence in Jones et al. (2009), suggesting possible conflicting signals in the sequence. Our analyses agree with the findings of Weng et al. (2012) and place *P. nanum* sister to the rest of clade A2, albeit with poor support. This finding is in conflict with other studies (Bakker et al., 2004; Röschenbleck et al., 2014) that used plastome indel coding as well as rDNA ITS sequence data. Therefore, confirmation from additional nuclear genomic data is needed.

The placement of *P. karooicum* within section *Subsucculentia* based on *rbcL* data and chromosome number ($x = 10$) by van der Walt et al. (1995) has been problematic. Section *Subsucculentia* species were

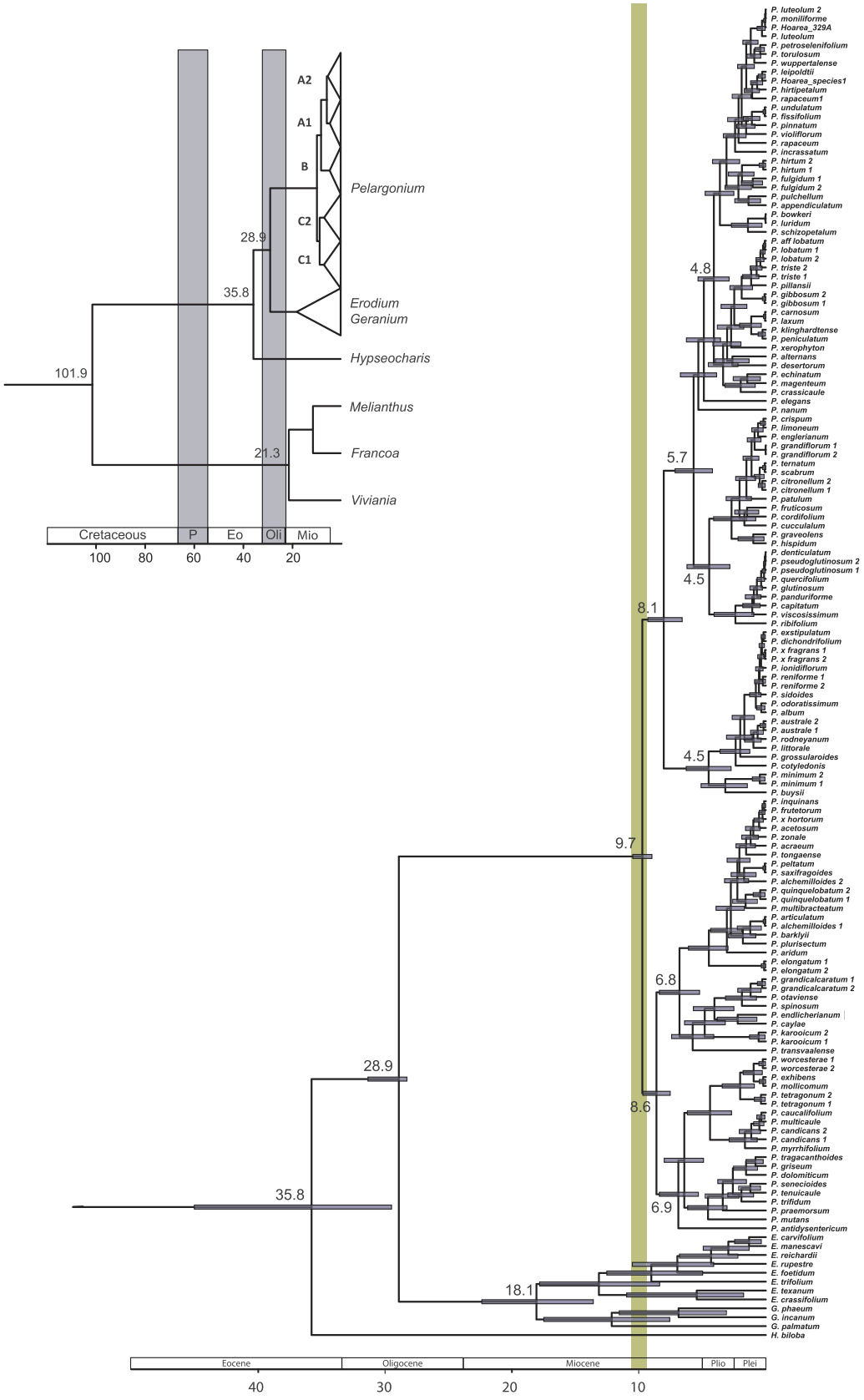

Fig. 4. (A) Time-calibrated phylogenetic tree of *Pelargonium*. Horizontal bars represent 95% highest posterior density (HPD) around mean node ages. Green line indicates Miocene-Pliocene climate change, used for ecological calibration (see text). (B) Cartoon style phylogenetic tree of *Pelargonium* showing outgroups used (dates are based on this study). Capital letters correspond to main clades: V = *Viviania*, F = *Francoa*, M = *Melianthus*, H = *Hypseocharis biloba*, G = *Geranium*, E = *Erodium*, P = *Pelargonium*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

previously considered monophyletic based on a shared base chromosome number of $x = 10$ (van der Walt et al., 1995). This was in conflict with findings from previous cpDNA based phylogenetic studies that indicated *P. karooicum* ($x = 10$) is part of a clade including *P. quercetorum* ($x = 17$), *P. endlicherianum* ($x = 17$) and *P. caylae* ($x = 9$) (Bakker et al., 2004, 2000a; Röschenbleck et al., 2014) Our analyses place *P. karooicum* as sister to a small clade formed by these $x = 9$, 10 17 species, still making the $x = 10$ species paraphyletic. This unresolved placement, the shared base chromosome number with sect. *Subsucculentia* but similar morphology to sect. *Jenkinsonia* species and multiple ribotypes (Bakker, unpubl. data), suggests that *P. karooicum* might be the result of an ancient hybridisation event (Röschenbleck et al., 2014; van der Walt et al., 1995).

In contrast to Röschenbleck et al. (2014), sections *Ligularia* and *Hoarea* are each monophyletic and we recovered increased resolution for the section *Ligularia*. This finding is important as the evolution of the formation of tunicate tubers coupled with a geophytic growth form in sect. *Hoarea* can now be studied in a proper sister-group context.

Although the existence of a *Polyactium-Otidia-Cortusina* clade has been disputed by Röschenbleck et al. (2014) and Weng et al. (2012), we find a highly supported (100%) *POC* clade that also includes *P. desertorum, P. alternans* and *P. xerophyton.* The latter three were designated 'unplaced' taxonomically within subgenus *Pelargonium* by Röschenbleck et al. (2014). *Pelargonium desertorum* as well as *P. xerophyton* have previously been assigned to section *Cortusina sensu stricto* (based on vegetative characters, Dreyer et al., 1992; Röschenbleck et al., 2014) and *P. alternans* for a long time has been part of the sect. *Otidia* based on its succulent stems (Röschenbleck et al., 2014). Rather than leaving them 'unplaced', we suggest restoring these species to their respective sections taxonomically and affirm the *POC* clade including *P. desertorum*, *P. xerophyton* and *P. alternans.* Upon further character and taxon sampling, it is possible that these species will be resolved in their respective *Otidia* and *Cortusina* clades.

Section *Campylia,* here represented by only *P. elegans,* appears to be sister to clade A2, which is in line with findings by Röschenbleck et al. (2014). However, as with *P. nanum* and its placement sister to clade A1, support for this finding is surprisingly weak. Again, data from different genomic compartments may help to clarify the phylogeny of this section. In addition, the inclusion of remaining species from the section is desirable since this will help to resolve the phylogenetic placement of the species in this section.

As in Röschenbleck et al. (2014), we retrieved the species *P. transvaalense*, *P. caylae*, *P. endlicherianum* and *P. karooicum* taxonomically unplaced as sister species to a clade formed by section *Subsucculentia*. We realize that phylogenetic patterns alone may be insufficient evidence to change existing taxonomic opinion and that corroboration from morphology and other evidence is necessary. With all the resources now at our disposal, it would be desirable to develop and classify all known species in the genus *Pelargonium*, i.e. to avoid having unplaced species. This would mean having a broader concept for groups such as sect. *Subsucculentia*.

4.2. Pelargonium dating

In our dating analysis, the influence of the 2nd calibration method on estimation of dates is apparent (Fig. 3). In the analysis for the calibration methods separately, the age estimates resulting from this method are much older compared with the Fossil and Ecological calibration methods. In combination with Fossil evidence, there still is a heavy influence of the 2nd calibration method visible in the resulting dates. We consider these results with some hesitancy because of the known problems with dating analyses based on solely 2nd calibration (Schenk, 2016), such as "false impression of precision" and "age estimates shifting away from those based on primary calibration".

Compared with the influence of 2nd calibration, the influence of Fossil evidence on date estimates is much less evident. We expected larger uncertainty in the Fossil based age estimates since all available fossils correspond to clades that are rather distantly related to *Pelargonium*. For example, since the Vivianaceae fossil dated at ~10 Mya is on a relatively long branch from *Pelargonium* it could be expected to introduce considerable dating uncertainty.

The inclusion of ecological calibration (based on climatic data) seems to have a much larger influence on the age estimates. In the separate analysis, the results are in the same range as the Fossil based results. The combination with 2nd calibration caused the (otherwise much older date estimates) to be dramatically lowered.

Based on the total evidence scenario (which we prefer since it is most inclusive, Fig. 4), our findings are similar to Fiz et al. (2008) and Palazzesi et al. (2012) who estimated an age of 10–15 My old for the *Pelargonium* crown node age based on pollen fossils, but have lower estimated node ages than in Bakker et al. (2005) and Verboom et al. (2009). The latter estimated *Pelargonium* crown to be approximately 30–35 Mya, and *Pelargonium* was considered to be older than most CFR lineages included in that study. Our findings, however, indicate the *Pelargonium* crown node originated around 9.7 Mya, which would be consistent with the average age of Fynbos lineages of 8.5 ± 1.85 Mya, and that of Succulent Karoo lineages of 5.17 ± 0.64 Mya as inferred across CFR clades by Verboom et al. (2009). In our study the Xerophytic clade A2 crown node, harbouring many Succulent Karoo species, was dated 5.3 Mya, consistent with the radiations found for other typical Cape Floristic Region clades (Bouchenak-Khelladi and Linder, 2017; Hughes et al., 2015; Linder, 2008, 2003; Linder and Verboom, 2015). As suggested previously, the pattern of nested radiations in *Pelargonium* Winter-rainfall region clade A2 could be the result of a radiation in response to aridification in the mid-Miocene, in addition to the ensuing fragmentation of niches, and could be an explanation for the high number of growth forms found in *Pelargonium* (Bakker et al., 2005; Verboom et al. 2009).

These finding shed new light on the remarkable biogeographic disjunctions in Cape – non Cape sister species distributions found in *Pelargonium*. Several *Pelargonium* species, especially from clade C, occur in high-altitude East African regions, extending to Ethiopia and Asia Minor, and stemming from Eastern Cape affinities (Bakker et al. 2005). These splits with such divergent distributions have all become much more recent compared with findings of Bakker et al. (2005). For example, the disjunction of *P. karooicum* (Cape) – *P. caylae* (Madagascar) – *P. endlicherianum* (Asia Minor) has now become as recent as ~5 Mya (early Pliocene). This and other occurrences of *Pelargonium* species outside the Greater Cape Floristic Region could be consistent with a 'Cape to Cairo' scenario as hypothesised for *Erica*, grasses and other clades in which the East African mountain range (starting from the Drakensbergen) provides a corridor across the equator (Galley et al., 2007). Whether the ancestral area for *Pelargonium* would have been inside or outside of the CFR remains unsolved.

5. Conclusions

Pelargonium phylogenetic relationships were estimated using a plastome-based data set including 74 plastid genes as well as the nuclear ITS region for 120 *Pelargonium* species, covering 43% of known species and 100% of known main clades. All species were retrieved within their expected major clade, i.e. consistent with previous phylogenetic studies. Resolution within clades has been increased compared to the last and most-inclusive study by Röschenbleck et al. (2014). We used different calibration approaches that have so far not been combined in one dating analysis yielding a crown node age for *Pelargonium* of 9.7 My, a much younger than previously expected. We present an improved, time-calibrated, phylogenetic framework for *Pelargonium* that can serve a diverse array of future studies. In particular we find the *Pelargonium* crown clade to be significantly younger than previously estimated, which makes it 'fit in' hypotheses of Fynbos and Succulent Karoo evolution much better. In order to arrive at a monophyletic

section-level classification more sequence data from additional genomic compartments is needed. Ideally, a combination of population sampling and multispecies coalescent analysis (Kubatko and Degnan, 2007) yielding formal species trees would form the basis for such a classification.

Acknowledgements

The Bakker lab thanks Di Lei for her contribution to the DNA labwork and Verena Berchtold from the Botanical Garden in Zürich. C. D. Schlichting and C. S. Jones thank Cape Nature, the Northern Cape Province and the Eastern Cape of South Africa for permission to collect leaf samples. We also thank the Department of Botany, University of Cape Town, and especially Edwina Marinus at the Compton Herbarium, Kirstenbosch for logistical support. We thank Kerri Mocko, Hugo Martinez-Cabrera, Timothy Moore and Caroline Chong for their time spent collecting and preparing samples and voucher specimens. Support for this project was provided by a US National Science Foundation funded Dimensions of Biodiversity grant (DEB-1046328) to C. D. S, C. S. J. and colleagues Kent Holsinger and John Silander. The Jansen lab thanks TACC at the University of Texas at Austin for access to supercomputers, Robin Parer from Geranicaeae.com for providing live plants of selected species and Plant Resources Center at the University of Texas at Austin for storage of voucher specimens. Support for this project was provided by a National Science Foundation grant (IOS-1027259) to R.K.J. and T.A.R. and the S. F. Blake Centennial Professorship to R.K.J.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.03.021.

References

Albers, F., van der Walt, J.J.A., 2007. Geraniaceae. In: The Families and Genera of Vascular Plants. pp. 157–167.

Ali, R.H., Bark, M., Miró, J., Muhammad, S.A., Sjöstrand, J., Zubair, S.M., Abbas, R.M., Arvestad, L., 2017. VMCMC: a graphical and statistical analysis tool for Markov chain Monte Carlo traces. BMC Bioinf. 18, 1–8. https://doi.org/10.1186/s12859-017-1505-3.

Bakker, F.T., Breman, F., Merckx, V., 2006. DNA sequence evolution in fast evolving mitochondrial DNA *nad*1 exons in Geraniaceae and Plantaginaceae. Taxon 55, 887–896.

Bakker, F.T., Culham, A., Daugherty, L.C., Gibby, M., 1999. A trnL-F based phylogeny for species of (Geraniaceae) with small chromosomes *Pelargonium*. Plant Species Biol. 216, 309–324.

Bakker, F.T., Culham, A., Gomez-Martinez, R., Carvalho, J., Compton, J., Dawtrey, R., Gibby, M., 2000a. Patterns of nucleotide substitution in angiosperm cpDNA *trnL* (UAA)-*trnF* (GAA) regions. Mol. Biol. Evol. 17, 1146–1155.

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

Bakker, F.T., Culham, A., Marais, E.M., Gibby, M., 2005. Nested radiation in Cape Pelargonium, in: Plant species-level systematic: new perspectives on pattern & process, pp. 75–100.

Bakker, F.T., Culham, A., Pankhurst, C.E., Gibby, M., 2000b. Mitochondrial and chlorplast DNA-based phylogeney of *Pelargonium* (Geraniaceae). Am. J. Bot. 87, 727–734.

Bakker, F.T., Hellbrügge, D., Culham, A., Gibby, M., 1998. Phylogenetic relationships within *Pelargonium* sect. *Peristera* (Geraniaceae) inferred from nrDNA and cpDNA sequence comparisons. Plant Syst. Evol. 211, 273–287.

Bakker, F.T., Lei, D., Yu, J., Mohammadin, S., Wei, Z., van de Kerke, S., Gravendeel, B., Nieuwenhuis, M., Staats, M., Alquezar-Planas, D.E., Holmer, R., 2016. Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organanelle Genome Assembly pipeline. Biol. J. Linn. Soc. 117, 33–43.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Becker, M., Albers, F., 2010. Pollinator shift and speciation in *Pelargonium alternans* (Geraniaceae). Schumannia 6, 207–218.

Becker, M., Albers, F., 2009. Taxonomy and phylogeny of two subgroups of *Pelargonium* section *Otidia* (Geraniaceae). 1. The *Pelargonium carnosum* complex. Bothalia 39, 73–85.

Blazier, C.C., Guisinger, M.M., Jansen, R.K., 2011. Recent loss of plastid-encoded ndh genes within *Erodium* (Geraniaceae). Plant Mol. Biol. 76, 263–272.

Blazier, J.C., Jansen, R.K., Mower, J.P., Govindu, M., Zhang, J., Weng, M.-L., Ruhlman, T. A., 2016a. Variable presence of the inverted repeat and plastome stability in Erodium. Ann. Bot. doi: 10.1093/aob/mcw065.

Blazier, J.C., Ruhlman, T.A., Weng, M.-L., Rehman, S.K., Sabir, J.S.M., Jansen, R.K., 2016b. Divergence of RNA polymerase α subunits in angiosperm plastid genomes is mediated by genomic rearrangement. Sci. Rep. 6, 24595.

Blerot, B., Baudino, S., Prunier, C., Demarne, F., Toulemonde, B., Caissard, J.-C., 2015. Botany, agronomy and biotechnology of Pelargonium used for essential oil production. Phytochem. Rev. 15, 935–960.

Bouchenak-Khelladi, Y., Linder, H.P., 2017. Frequent and parallel habitat transitions as driver of unbounded radiations in the Cape flora. Evolution (N.Y) 71, 2548–2561. https://doi.org/10.1111/evo.13364.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Comput. Biol. 10, 1–6. https://doi.org/10.1371/journal.pcbi.1003537.

Chumley, T.W., Palmer, J.D., Mower, J.P., Fourcade, M.H., Calie, P.J., Boore, J.L., Jansen, R.K., 2006. The complete chloroplast genome sequence of *Pelargonium x hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol. Biol. Evol. 23, 2175–2190.

Clark, S.C., Egan, R., Frazier, P.I., Wang, Z., 2013. ALE: a generic assembly likelihood evaluation framework for assessing the accuracy of genome and metagenome assemblies. Bioinformatics 29, 435–443. https://doi.org/10.1093/bioinformatics/bts723.

De Candolle, A.P., 1824. Geraniaceae. Prodromus systematis naturalis regni vegetabilis, vol. 1 Treutel, 637–682.

Doyle, J., 1991. DNA protocols for plants. Mol. Tech. Taxon. 283–293.

Dreyer, L.L., Albers, F., Van der Walt, J.J.A., Marschewski, D.E., 1992. Subdivision of *Pelargonium* sect. *Cortusina* (Geraniaceae). Plant Syst. Evol. 183, 83–97.

Drummond, A., Suchard, M., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Ferreira, M.A.R., Suchard, M.A., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Can. J. Stat. 36, 355–368. https://doi.org/10.1002/cjs.5550360302.

Fiz, O., Vargas, P., Alarcón, M., Aedo, C., García, J.L., Aldasoro, J.J., 2008. Phylogeny and historical biogeography of geraniaceae in relation to climate changes and pollination ecology phylogeny. Syst. Bot. 33, 326–342.

Galley, C., Bytebier, B., Bellstedt, D.U., Peter Linder, H., 2007. The Cape element in the afrotemperate flora: from Cape to Cairo? Proc. R. Soc. B Biol. Sci. 274, 535–543. https://doi.org/10.1098/rspb.2006.0046.

Goldblatt, P., Savolainen, V., Porteous, O., Sostaric, I., Powell, M., Reeves, G., Manning, J.C., Barraclough, T.G., Chase, M.W., 2002. Radiation in the Cape flora and the phylogeny of peacock irises Moraea (Iridaceae) based on four plastid DNA regions. Mol. Phylogenet. Evol. 25, 341–360. https://doi.org/10.1016/S1055-7903(02)00235-X.

Guisinger, M.M., Kuehl, J.V., Boore, J.L., Jansen, R.K., 2011. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. Mol. Biol. Evol. 28, 583–600.

Guisinger, M.M., Kuehl, J.V., Boore, J.L., Jansen, R.K., 2008. Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. Proc. Natl. Acad. Sci. USA 105, 18424–18429.

Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58, 367–380. https://doi.org/10.1093/sysbio/syp035.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

Hughes, C.E., Nyffeler, R., Linder, H.P., 2015. Evolutionary plant radiations: where, when, why and how? New Phytol. 207, 249–253. https://doi.org/10.1111/nph.13523.

James, C.M., Gibby, M., Barrett, J.A., 2004. Molecular studies in *Pelargonium* (Geraniaceae). A taxonomic appraisal of section *Ciconium* and the origin of the "Zonal" and "Ivy-leaved" cultivars. Plant Syst. Evol. 243, 131–146.

Jones, C.S., Bakker, F.T., Schlichting, C.D., Nicotra, A.B., 2009. Leaf Shape Evolution in the South African Genus Pelargonium L' Hér. (Geraniaceae). Evolution (N. Y). 63, 479–497.

Jones, C.S., Cardon, Z.G., Czaja, A., 2003. A Phylogenetic view of low-level CAM in *Pelargonium* (Geraniaceae). Am. J. Bot. 90, 135–142.

Jones, C.S., Price, R.A., 1995. Diversity and evolution of seedling Bauplane in *Pelargonium* (Geraniaceae). Aliso 4, 281–295.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol.

Biol. Evol. 29, 1695–1701.
Linder, H.P., 2008. Plant species radiations: Where, when, why? Philos. Trans. R. Soc. B Biol. Sci. 363, 3097–3105. https://doi.org/10.1098/rstb.2008.0075.
Linder, H.P., 2003. The radiation of the Cape flora, southern Africa. Biol. Rev. 78, 597–638.
Linder, H.P., Verboom, G.A., 2015. The evolution of regional species richness: the history of the Southern African Flora. Annu. Rev. Ecol. Evol. Syst. 46, 393–412. https://doi.org/10.1146/annurev-ecolsys-112414-054322.
Liu, L., Yu, L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. Mol. Phylogenet. Evol. 53, 320–328. https://doi.org/10.1016/j.ympev.2009.05.033.
Maddison, W.P., Maddison, D.R., 2015. Mesquite: a modular system for evolutionary analysis. Version 3, 04.
Manning, J., Goldblatt, P., 2012. Plants of the greater cape floristic region 1: the core cape flora. Strelitzia 29, 1–870. https://doi.org/10.1017/CBO9781107415324.004.
Manning, J.C., Euston-Brown, D.I.W., Magee, A.R., 2015. *Pelargonium uliginosum* (Geraniaceae: Section *Hoarea*), a new species from Western Cape, South Africa, and an updated key to the species of the *P. dipetalum* group. South African J. Bot. 97, 204–207. https://doi.org/10.1016/j.sajb.2015.01.016.
Marais, E.M., 2016. Five new species of *Pelargonium*, section *Hoarea* (Geraniaceae), from the Western and Northern Cape Provinces of South Africa. South African J. Bot. 103, 145–155.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. Proc. Gatew. Comput. Environ. Work. 1–8.
Moore, T.E., Schlichting, C.D., Aiello-Lammens, M.E., Mocko, K., Jones, C.S., 2018. Divergent trait and environment relationships among parallel radiations in *Pelargonium* (Geraniaceae): a role for evolutionary legacy? New Phytol. https://doi.org/10.1111/nph.15196.
Mower, J.P., Touzet, P., Gummow, J.S., Delph, L.F., Palmer, J.D., 2007. Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants. BMC Evol. Biol. 7.
Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274. https://doi.org/10.1093/molbev/msu300.
Nicotra, A.B., Chong, C., Bragg, J.G., Ong, C.R., Aitken, N.C., Chuah, A., Lepschi, B., Borevitz, J.O., 2016. Population and phylogenomic decomposition via genotyping-by-sequencing in Australian *Pelargonium*. Mol. Ecol. 25, 2000–2014. https://doi.org/10.1111/mec.13584.
Nicotra, A.B., Cosgrove, M.J., Cowling, A., Schlichting, C.D., Jones, C.S., 2008. Leaf shape linked to photosynthetic rates and temperature optima in South African *Pelargonium* species. Oecologia 154, 625–635.
Palazzesi, L., Gottschling, M., Barreda, V., Weigend, M., 2012. First Miocene fossils of Vivianiaceae shed new light on phylogeny, divergence times, and historical biogeography of Geraniales. Biol. J. Linn. Soc. 107, 67–85.
Park, S., Grewe, F., Zhu, A., Ruhlman, T.A., Sabir, J., Mower, J.P., Jansen, R.K., 2015. Dynamic evolution of *Geranium* mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol. 208, 570–583.
Parkinson, C.L., Mower, J.P., Qiu, Y.-L., Shirk, A.J., Song, K., Young, N.D., DePamphilis, C.W., Palmer, J.D., 2005. Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol, Biol, pp. 5.
Price, R.A., Palmer, J.D., 1993. Phylogenetic relationships of the geraniaceae and geraniales from *rbcL* sequence comparisons. Ann. Missouri Bot. Gard. 80, 661–671.
Promega, 2010. Wizard® DNA Clean-Up System.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. https://doi.org/10.1093/sysbio/sys029.
Röschenbleck, J., Albers, F., Müller, K., Weinl, S., Kudla, J., 2014. Phylogenetics, character evolution and a subgeneric revision of the genus *Pelargonium* (Geraniaceae). Phytotaxa 159, 31–76.
Ruhlman, T.A., Jansen, R.K., 2018. Aberration or analogy? The atypical plastomes of Geraniaceae, in: Chaw, S.-M., Jansen, R.K. (Eds.), Advances in Botanical Research 85: Plastid Genome Evolution. Elsevier, Amsterdam, pp. 223–262. doi: https://doi.org/10.1016/bs.abr.2017.11.017.
Schenk, J.J., 2016. Consequences of secondary calibrations on divergence time estimates. PLoS ONE 11. https://doi.org/10.1371/journal.pone.0148228.
Snijman, D.A. (Ed.), 2013. Plants of the Greater Cape Floristic Region 2: The Extra Cape Flora. Strelitzia 30, 1–557.
Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stöver, B.C., Müller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinf. 11, 1.
Sweet, R., 1822. Geraniaceae 1. Ridgway, J. 1–100.
Swofford, D.L., 2002. Phylogenetic analysis using parsimony (*and Other Methods). Version 4.
Thermo Fisher Scientific, D., 2008. Nanodrop.
Touloumenidou, T., Bakker, F.T., Marais, E.M., Albers, F., 2003. Chromosomal evolution interpreted from the rDNA ITS phylogeny for *Pelargonium* sect. *Hoarea (Geraniaceae)*. Schumannia 4, 1–14.
Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., Minh, B.Q., 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235. https://doi.org/10.1093/nar/gkw256.
Vaidya, G., Lohman, D.J., Meier, R., 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27, 171–180.
van der Walt, J.J.A., 1985. A taxonomic revision of the type section of *Pelargonium* L ' Hérit. Both 15, 345–385.
Van der Walt, J.J.A., Albers, F., Gibby, M., 1990. Delimitation of *Pelargonium* sect. *Glaucophyllum* (Geraniaceae). Plant Syst. Evol. 171, 15–26.
van der Walt, J.J.A., Albers, F., Gibby, M., Marschewski, D.E., Price, R.A., 1995. A biosystematic study of *Pelargonium* section *Ligularia*: 1. A new section *Subsucculentia*. South African J. Bot. 61, 339–346.
Van der Walt, J.J.A., Vorster, P.J., 1988. Pelargoniums of Southern Africa. Volumes I-III. Annals of Kirtenbosch Botanic Gardens.
Van der Walt, J.J.A., Vorster, P.J., 1983. Phytogeography of *Pelargonium*. Bothalia 14, 517–523.
Van der Walt, J.J.A., Vorster, P.J., 1981. Typification of the genus Pelargonium L'Herit. (fam. Geraniaceae). Taxon 30, 307.
Verboom, G.A., Archibald, J.K., Bakker, F.T., Bellstedt, D.U., Conrad, F., Dreyer, L.L., Forest, F., Galley, C., Goldblatt, P., Henning, J.F., Mummenhoff, K., Linder, H.P., Muasya, A.M., Oberlander, K.C., Savolainen, V., Snijman, D., van der Niet, T., Nowell, T.L., 2009. Origin and diversification of the Greater Cape flora: ancient species repository, hot-bed of recent radiation, or both? Mol. Phylogenet. Evol. 51, 44–53. https://doi.org/10.1016/j.ympev.2008.01.037.
Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., Soltis, D.E., 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. 106, 3853–3858.
Warren, D.L., Geneva, A.J., Lanfear, R., 2017. RWTY (R We There Yet): an R package for examining convergence of Bayesian phylogenetic analyses. Mol. Biol. Evol. 34 (4), 1016–1020.
Weng, M.-L., Blazier, J.C., Govindu, M., Jansen, R.K., 2013. Reconstruction of the ancestral plastid genome in geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. Mol. Biol. Evol. 31, 645–659.
Weng, M.-L., Ruhlman, T.A., Gibby, M., Jansen, R.K., 2012. Phylogeny, rate variation, and genome size evolution of *Pelargonium* (Geraniaceae). Mol. Phylogenet. Evol. 64, 654–670. https://doi.org/10.1016/j.ympev.2012.05.026.
White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR Protoc. 18, 315–322. https://doi.org/10.1016/B978-0-12-372180-8.50042-1.
Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., Li, S., Zhou, X., Lam, T.W., Li, Y., Xu, X., Wong, G.K.S., Wang, J., 2014. SOAPdenovo-Trans: De novo transcriptome assembly with short RNA-Seq reads. Bioinformatics 30, 1660–1666.