RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

A phylogenetic study of the tribe Antirrhineae: Genome duplications and long-distance dispersals from the Old World to the New World[1]

Ezgi Ogutcen[2] and Jana C. Vamosi

PREMISE OF THE STUDY: Antirrhineae is a large tribe within Plantaginaceae. Mostly concentrated in the Mediterranean Basin, the tribe members are present both in the Old World and the New World. Current Antirrhineae phylogenies have different views on taxonomic relationships, and they lack homogeneity in terms of geographic distribution and ploidy levels. This study aims to investigate the changes in the chromosome numbers along with dispersal routes as definitive characters identifying clades.

METHODS: With the use of multiple DNA regions and taxon sampling enriched with de novo sequences, we provide an extensive phylogeny for Antirrhineae. The reconstructed phylogeny was then used to investigate changes in ploidy levels and dispersal patterns in the tribe using ChromEvol and RASP, respectively.

KEY RESULTS: Antirrhineae is a monophyletic group with six highly supported clades. ChromEvol analysis suggests the ancestral haploid chromosome number for the tribe is six, and that the tribe has experienced several duplications and gain events. The Mediterranean Basin was estimated to be the origin for the tribe with four long-distance dispersals from the Old World to the New World, three of which were associated with genome duplications.

CONCLUSIONS: On an updated Antirrhineae phylogeny, we showed that the three out of four dispersals from the Old World to the New World were coupled with changes in ploidy levels. The observed patterns suggest that increases in ploidy levels may facilitate dispersing into new environments.

KEY WORDS Antirrhineae; genome duplication; geographical origin; long-distance dispersal; phylogeny; Plantaginaceae

For the last two decades, advances in molecular techniques and fast sequencing methods have increased the number of studies utilizing molecular markers in phylogenetic reconstruction of a wide range of taxa (Peters et al., 2011). Switching from morphological markers to molecular markers has caused some major changes in the circumscription of many taxonomic groups. Formerly part of the family Scrophulariaceae Juss., Antirrhineae Dumort. is now considered one of the 12 tribes under Plantaginaceae Juss. (Albach et al., 2005). An extensive revision of the tribe recognizes 27 genera representing over 300 species within Antirrhineae (Sutton, 1988). There is considerable morphological variation in the tribe in terms of floral shape and color (Ghebrehiwet et al., 2000), but the clades do not have clear synapomorphies. An earlier study found one synapomorphic character (isodiametric leaf shape) for the *Maurandya* clade, and two (medium hilum position and ring-forming anthers) for the *Anarrhinum* clade (Ghebrehiwet et al., 2000). Other widely studied characters such as floral color, habit, presence of spurs, and life span seem to be labile in the tribe (Sutton, 1988). The diverse floral characteristics foster the adoption of some members of the tribe as popular garden flowers including snapdragons (*Antirrhinum*) and toadflax (*Linaria*).

Previous studies all supported the monophyly of the tribe Antirrhineae, but there has not been agreement on the subtribal relationships (Ghebrehiwet et al., 2000; Vargas et al., 2004; Fernández-Mazuecos et al., 2013; Guzman et al., 2015). Based on earlier morphological studies (Rothmaler, 1943; Rothmaler, 1956), Antirrhineae were divided into five subtribes: Anarrhinae (*Anarrhinum*), Gambeliinae (*Galvezia*, *Sacculariа*, and *Gambelia*), Linariinae (*Asarina, Cymbalaria, Kickxia, Linaria, Chaenorrhinum, Antirrhinum, Neogaerrhinum, Acanthorrhinum, Schweinfurthia, Pseudorontium, Misopates*), Maurandyinae (*Rhodochiton, Lophospermum, Epixiphium, Maurandya, Maurandella*), and Mohaveinae (*Mohavea*).

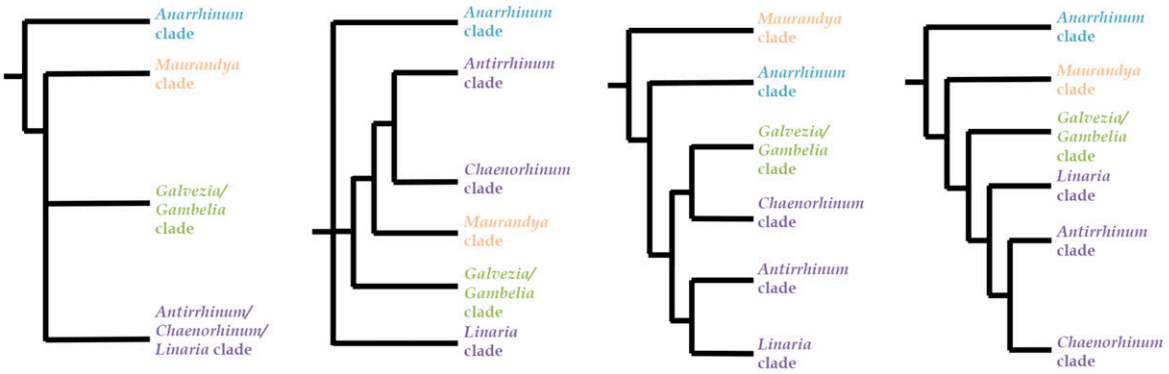
[1] Manuscript received 3 November 2015; revision accepted 14 April 2016.
Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada
[2] Author for correspondence (e-mail: eogutcen@ucalgary.ca)
doi:10.3732/ajb.1500464

More recent studies based on molecular data proposed different subtribal classifications (Fig. 1). The Antirrhineae phylogeny reconstructed with the plastid marker *ndhF* (Ghebrehiwet et al., 2000) divided the tribe into four clades: *Anarrhinum* clade, *Antirrhinum* clade, *Gambelia* clade, and *Maurandya* clade. Later phylogenies based on the nuclear marker *ITS* (Vargas et al., 2004; Fernández-Mazuecos et al., 2013) divided the previously named "*Antirrhinum* clade" into three groups: *Antirrhinum* group, *Chaenorhinum* group, and *Linaria* group. They also added more taxa to the phylogeny and had some differences in the placement of some genera. The most recent study used two DNA regions (*ITS* and *ndhF*) and had an extensive species sampling (Guzman et al., 2015). When compared to the earlier phylogenies (Ghebrehiwet et al., 2000; Vargas et al., 2004; Fernández-Mazuecos et al., 2013) this updated Antirrhineae phylogeny had major differences in the relationships among the clades.

In terms of geographical distribution, the previously published Antirrhineae clades lack homogeneity: *Anarrhinum, Chaenorhinum*, and *Linaria* clades are composed of Old World genera; *Antirrhinum, Galvezia/Gambelia*, and *Maurandya* clades have both Old World and New World representatives. The Old World taxa have a dense distribution in the Mediterranean Basin, but there are also some species found in other parts of Africa, Europe, and Asia. The New World taxa are mostly concentrated in the California coast, which has a Mediterranean climate, but there are also representatives of the tribe in other parts of North America and South America (Hassler, 2016). An Old World origin for the tribe was previously suggested (Vargas et al., 2014), but the study only involved a single representative species from each genus.

Polyploidy is defined as having more than two copies of each chromosome, and it is a common feature in plants (Wood et al., 2009). It was proposed that most of the genetic material found in angiosperms today is provided by the ancient polyploidization events that occurred during early angiosperm evolution (De Bodt et al., 2005). Based on the number of chromosome sets in extant species, it is estimated that up to 100% of all the angiosperms have undergone chromosome number changes throughout their evolution (Wood et al., 2009). In addition to genome duplication, where the haploid chromosome number is doubled (or partially duplicated), changes in the chromosome numbers can also be achieved by dysploidy, where single chromosomes, instead of the full genome, are gained or lost (Mayrose et al., 2010). These chromosomal changes were shown to be commonly observed throughout angiosperm evolution (Escudero et al., 2014).

Recently, chromosome number has been investigated in conjunction with distribution as an important synapomorphy characterizing clades (Glick and Mayrose, 2014; Linder and Barker, 2014). The emerging association between long-distance dispersal events



	Ghebrehiwet *et al.*, 2000	Vargas *et al.*, 2004	Fernández-Mazuecos *et al.*, 2013	Guzman *et al.*, 2015
***Maurandya* clade**	*Asarina, Cymbalaria, Maurandella, Maurandya*, Rhodochiton*	*Asarina, Cymbalaria, Lophospermum*, Maurandella, Rhodochiton*	*Asarina, Cymbalaria, Epixiphium*, Holmgrenanthe*, Lophospermum*, Mabrya*, Maurandella, Maurandya*, Rhodochiton*	*Asarina, Cymbalaria, Epixiphium*, Holmgrenanthe*, Lophospermum*, Mabrya*, Maurandella, Maurandya*, Rhodochiton*
***Antirrhinum* clade**	*Antirrhinum, Chaenorhinum, Howelliella, Linaria, Misopates, Mohavea, Schweinfurthia*	*Acanthorrhinum*, Antirrhinum, Galvezia, Howelliella, Misopates, Mohavea*, Neogaerrhinum*, Pseudomisopates*, Sairocarpus**	*Acanthorrhinum*, Antirrhinum, Gambelia, Howelliella, Misopates, Mohavea*, Neogaerrhinum*, Pseudomisopates*, Sairocarpus**	*Acanthorrhinum*, Antirrhinum, Gambelia, Howelliella, Misopates, Mohavea*, Neogaerrhinum*, Pseudomisopates*, Sairocarpus**
***Chaenorhinum* clade**		*Albraunia*, Chaenorhinum, Holzneria**	*Albraunia*, Chaenorhinum, Holzneria**	*Albraunia*, Chaenorhinum, Holzneria**
***Linaria* clade**		*Linaria*	*Linaria*	*Linaria*
***Galvezia/Gambelia* clade**	*Galvezia, Gambelia**	*Pseudorontium*, Schweinfurthia*	*Galvezia, Pseudorontium*, Schweinfurthia*	*Galvezia, Pseudorontium*, Schweinfurthia*
***Anarrhinum* clade**	*Anarrhinum, Kickxia*	*Anarrhinum, Kickxia*	*Anarrhinum, Kickxia*	*Anarrhinum, Kickxia, Nanorrhinum**

FIGURE 1 Schematic representations and the circumscriptions of the previously published Antirrhineae phylogenies. Trees are modified from Ghebrehiwet et al., 2000, Vargas et al., 2004, Fernández-Mazuecos et al., 2013, and Guzman et al., 2015. Clade names were changed to accommodate the differences in their circumscriptions in the papers. Underlined genera are placed under different clades in different phylogenies; genera with an asterisk (*) are missing in at least one study.

and genome duplications has the potential to connect pattern to process when we consider the ways in which polyploidization can expand the ecological range a given lineage can colonize (Brochmann et al., 2004; Vamosi and McEwen, 2013). Despite the traditional view that polyploidization is an evolutionary dead end (Stebbins, 1971; Mayrose et al., 2011), recent studies show poylploidy and diversification rates are correlated in some angiosperm taxa (de Bodt et al., 2005; Vamosi and Dickinson, 2006).

Covering a large number of species sampled from different genera, we reconstructed the phylogeny of Antirrhineae using both nuclear and plastid markers. We further aim to (1) investigate the monophyletic clades in Antirrhineae, (2) determine the changes in the haploid chromosome numbers throughout the phylogeny, (3) explore the geographic origin of the tribe, and (4) examine the dispersal routes throughout the diversification of the tribe.

MATERIALS AND METHODS

Taxon sampling—Sequence data were available in the National Center for Biotechnology Information (NCBI) GenBank database for 146 species from 24 out of 27 recognized genera within Antirrhineae (Sutton, 1988). We enriched the available sequence data with 19 de novo sequences for six species (*Chaenorhinum minus* (L.) Lange, *Kickxia lanigera* Hand.-Mazz., *K. spuria* (L.) Dumort., *Linaria dalmatica* (L.) Mill., *L. repens* Mill., *L. vulgaris* Mill; Appendix 1). Eleven outgroups were selected for the study, with one species representing each tribe under Plantaginaceae. Overall, the analysis included 157 species (Appendix S1; see Supplemental Data with the online version of this article).

Five DNA regions were used in the study. The *ITS* region was selected as the nuclear marker, and includes the 18S ribosomal RNA gene, internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2, and 26S ribosomal RNA gene (collectively called *ITS* from now on). The plastid markers selected were the following: nicotinamide adenine dinucleotide dehydrogenase F gene (*ndhF*); the gene encoding the large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (*rbcL*); ribosomal protein S16 gene (*rps16*); and the region that includes the transfer RNA-Leu gene and the intergenic spacer between the transfer RNA-Leu gene and transfer RNA-Phe (*trnL-F)*.

DNA extraction, amplification, and sequencing—Dried plant materials were collected from the University of Calgary Herbarium (Appendix 1). Total genomic DNA was extracted from the herbarium specimens using a modified CTAB protocol (Doyle and Doyle, 1987). The extracted DNA was used as template to PCR-amplify the selected markers using universal primers described in the literature (Taberlet et al., 1991; Baldwin, 1992; Oxelman et al., 1997; Dong et al., 2014; Appendix S2). The thermocycler protocol had an initial denaturation at 94°C for 2 min, followed by 30 cycles of amplification (94°C for 1 min, 45°C for 30 s, and 72°C for 1 min), followed by a final extension at 72°C for 4 min and a hold at 4°C. A 25 μl PCR reaction mixture contained 50–100 ng DNA, 2.5 units Taq DNA Polymerase with 10× buffer, 0.2 mM dNTP mix, 25 pmol primers, and 0.8 mg/ml BSA.

Amplified DNA markers were purified and prepared for DNA sequencing using a QiaQuick PCR Purification Kit (Qiagen, Valencia, California, USA). Samples to be sequenced were prepared with 50–100 ng sequencing template per 1000 base pairs, and 3.2 pmol of primer per sequencing reaction. The same primers were used for sequencing the amplified markers. The sequencing samples were submitted to the University of Calgary Genetic Analysis Laboratory for automated DNA sequencing. Sequencing reactions were performed using an Applied Biosystems 3730xl genetic analyzer (Applied Biosystems, Foster City, California, USA). The completed sequences were assembled and edited using Mesquite 3.02 (Maddison and Maddison, 2014) with the package Opal 2.1 (Wheeler and Kececioglu, 2007).

Sequence assembly, alignment and phylogenetic analysis—Both de novo sequences and the sequences collected from the NCBI GenBank database were combined and automatically aligned using Mesquite 3.02 with the package Opal 2.1. Sequences from five different markers were concatenated in two different combinations in Mesquite 3.02: (1) the combination of four plastid markers (*ndhF, rbcL, rps16,* and *trnL-F)* was named "4Pla," and (2) the combination of one nuclear marker and four plastid markers (*ITS, ndhF, rbcL, rps16,* and *trnL-F)* was named "5All". Although we approached with caution (see details in results), the incongruence length difference (ILD) test (Farris et al., 1994) was performed using PAUP* 4.0a147 (Swofford, 2003) to determine whether the markers were suitable for combining to obtain a large sequence data set for phylogenetic reconstruction of the tribe.

We used jModeltest (Guindon and Gascuel, 2003; Darriba et al., 2012) to select substitution models for the combined markers. The general time reversible model (Tavaré, 1986) with invariant sites and a gamma distribution of rates across sites (GTR+I+G) was selected using the Akaike Information Criterion (AIC; Akaike, 1974) and the Bayesian Information Criterion (BIC; Schwarz, 1978) as measures (both AIC and BIC selected the same model for all tests).

Maximum likelihood (ML) analyses were performed in RAxML 7.4.2 (Stamatakis, 2014) with both 4Pla and 5All marker combinations, and with multiple outgroups (see above for details). The GTR+I+G model was used with 1000 rapid bootstrap replicates. Trees were visualized and edited using FigTree 1.3.1 (Rambaut, 2009).

Bayesian analyses were performed using BEAST 1.8.1 (Drummond et al., 2012) with both 4Pla and 5All marker combinations. BEAST .xml files were generated using BEAUti 1.8.1 (Drummond et al., 2012). Multiple outgroups were selected (see above for details). A GTR+I+G model was used for both marker combinations. Relative node ages were estimated using an uncorrelated lognormal relaxed clock model with a normal mean distribution prior. For tree calibration, the divergence between *Chaenorrhinum* and *Linaria* with a mean of 23 million years ago (mya) and standard deviation of 4 million years (Myr) was used. This calibration was based on a previous study that incorporated available angiosperm fossil data (Bell et al., 2010). A pure Birth Process (Gernhard, 2008) was used as a tree prior. Two Markov Chain Monte Carlo (MCMC) chains were run for 10,000,000 generations and sampled every 1000 generations, resulting in 10,000 sampled trees per chain. Analyses were run until the effective sample size (ESS) values were above 200. Stationarity was verified using the program Tracer 1.5 (Rambaut et al., 2014). Maximum clade credibility (MCC) trees were generated using TreeAnnotator 1.8.1 (Heled and Bouckaert, 2013). The first 1000 (10%) tree samples from each chain were discarded as burn-in. The two tree data sets from the two MCMC chains were combined using LogCombiner v1.8.1 (Drummond et al., 2012). The generated trees were visualized using FigTree 1.3.1 (Rambaut, 2009). The MCC trees obtained from the Bayesian

analysis of both marker combinations (4Pla and 5All) were used for all the downstream analyses.

Chromosome count data and analysis of ploidy levels—Haploid chromosome counts for the Antirrhineae species were collected from the Chromosome Counts Database (Rice et al., 2015). In case of multiple chromosome counts for the same species, the most frequently sampled chromosome count was used. When the sampling frequencies were similar, the lowest chromosome count was used. Species with no chromosome count data were coded as ×.

We used ChromEvol software (Mayrose et al., 2010; Glick and Mayrose, 2014) to infer the ancestral haploid chromosome numbers in the tribe, and to determine at which points and how many times the ploidy levels have changed along the phylogeny. All possible models were tested and ranked based on their likelihood values (Akaike, 1974). "CONST_RATE_DEMI" model (the lowest AIC scoring model) was used in the analyses with constant rate for single chromosome increases (gain) and decreases (loss), and constant and equal rates for whole-genome duplication (polyploidy) and partial-genome duplication (demipolyploidy). Using both maximum likelihood and Bayesian approaches, 10,000 simulations were performed.

Dispersal-Vicariance analysis—We collected distribution data from the World Plants online database in Species 2000 & ITIS Catalogue of Life (Hassler, 2016). Some species were introduced into some regions, but only native distributions were included in the analyses. For the binary-state analysis, Old World (Asia, Europe, Africa) occurrence was denoted as A; New World occurrence (North America, South America) was denoted as B. For the multiple-state analysis, the distribution data were categorized as follows: Mediterranean region was denoted as A; Old World continents excluding the Mediterranean region as B; Americas as C.

We used RASP 2.1 beta software (Yu et al., 2015) to examine the dispersal-vicariance patterns and the ancestral distributions of the tribe. We used three different reconstruction methods: (1) Statistical Dispersal-Vicariance Analysis (Ronquist, 1997), (2) the Bayesian binary method (Ronquist, 2004), and (3) the Dispersal Extinction-Cladogenesis model (Ree and Smith, 2008). Default parameters were used for all the methods except for the Bayesian analysis; the cycle number was increased to 1,000,000, and estimated state frequencies (F81) with among-site variation with gamma distribution were used as the model.

RESULTS

Antirrhineae phylogeny—The nuclear marker *ITS* had sequence data for 149 species including the outgroups. After the alignment of all the species, the total length of the nuclear *ITS* region analyzed was 815 bp. For the plastid markers *ndhF*, *rbcL*, *rps16*, and *trnL-F*, we had sequence data for 106, 27, 23, and 76 species respectively. Once all the species were aligned, the analyzed plastid regions were 2124, 1408, 898, and 938 bp long respectively. When we aligned and combined all the plastid markers (4Pla) together, the resulting matrix contained 5368 characters, comprising 139 species including the outgroups. When all the markers (5All) were aligned and combined together, the resulting matrix was 6183 nucleotides long, with 157 species, including the outgroups.

In the Bayesian MCC tree constructed with the 5All sequence data set, Antirrhineae formed a monophyletic group (1.00 PP) with six well-supported clades (Fig. 2): The *Anarrhinum* clade (*Anarrhinum* Desf., *Kickxia* Dumort.; 0.93 PP), the *Maurandya* clade (*Asarina* Miller, *Cymbalaria* Hill, *Holmgrenanthe* Elisens, *Lophospermum* D.Don, *Mabrya* Elisens, *Maurandya* Ortega, *Rhodochiton* Zucc.; 0.97 PP), the *Gambelia* clade (*Gambelia* Nutt., *Galvezia* Dombey ex Juss., *Schweinfurthia* A.Braun; 0.96 PP), the *Chaenorhinum* clade (*Albraunia* Speta, *Chaenorhinum* (DC.) Reichenb., *Holzneria* Speta; 0.98 PP), the *Antirrhinum* clade (*Acanthorrrhinum* Rothm., *Antirrhinum* L., *Misopates* Rafin., *Mohavea* A.Gray, *Neogaerrhinum* Rothm., *Sairocarpus* D.A.Sutton; 0.99 PP), and the *Linaria* clade (*Linaria* Miller, *Nuttallanthus* D.A.Sutton; 0.98 PP). Overall, PPs were very high supporting the genera, but relatively lower at the species level, especially within large genera like *Antirrhinum* and *Linaria* (Fig. 2).

While the ILD is arguably sensitive and susceptible to Type I error (Barker and Lutzoni, 2002; Darlu and Lecointre, 2002), it detected no significant difference among the four plastid sequences. We thus combined our four plastid markers to use a larger sequence data set for phylogeny reconstruction (Appendix S3). According to the ILD test, there was a significant difference between the plastid sequence data set and the nuclear sequence data set, so we compared the reconstructed phylogenies that were generated using ITS only, 4Pla, and 5All sequence data sets. Phylogenetic reconstructions with ITS only, 4Pla, and 5All data sets produced similar tree topologies with a few differences. Overall, the phylogeny generated with the ITS-only data set had the lowest support values, and the topology was less similar to the phylogenies generated with the combined data sets (Appendix S4). The 4Pla and 5All data sets produced similar tree topologies except for one main difference: *Antirrhinum* was sister to the *Acanthorrhinum-Misopates* lineage in the 4Pla tree, whereas the sister for *Antirrhinum* was the *Sairocarpus-Mohavea* lineage in the 5All tree. When we partitioned the 5All data into two partitions (ITS and 4Pla), and compared the phylogeny generated using the partitioned data set with the phylogeny generated using the nonpartitioned data set, there were only minor differences at the nodes with low support values.

The ML and Bayesian analyses resulted in similar topologies for Antirrhineae, with the Bayesian MCC trees having slightly higher posterior probabilities (PP) than the ML bootstrap values (BS) near the tips. In the phylogenies generated using the 4Pla data set, *Schweinfurthia pterosperma* A. Braun was grouped together with the *Sairocarpus-Mohavea* lineage in the Bayesian analysis, whereas in the ML analysis, it was the outgroup of the *Linaria-Antirrhinum-Chaenorhinum* clade group. The relationship among the groups of the *Antirrhinum* clade also differed between the Bayesian and the ML analysis of the 4Pla data set. The Bayesian analysis showed the *Acanthorrhinum-Misopates* and the *Sairocarpus-Mohavea* lineages as sister taxa, whereas ML results place *Antirrhinum* and *Sairocarpus-Mohavea* group together as sister taxa. In the phylogenies generated using the 5All data set, there were two main differences between the Bayesian and ML analyses: (1) the *Gambelia* clade included *Gambelia speciosa* Nutt., *Schweinfurthia papilionacea* Boiss., and all three *Galvezia* species in the ML analysis (BS = 67%), whereas the Bayesian analysis placed *Galvezia juncea* (Benth.) Ball within the *Antirrhinum* clade, grouped with the *Sairocarpus-Mohavea* lineage (PP = 0.74); and (2) both *Neogaerrhinum* species were placed under the *Sairocarpus-Mohavea* lineage in the Bayesian analysis (PP < 0.50), whereas in the ML analysis, *Neogaerrhinum strictum* (Hook. & Arn.) Rothm. was the sister species of *Schweinfurthia pterosperma*

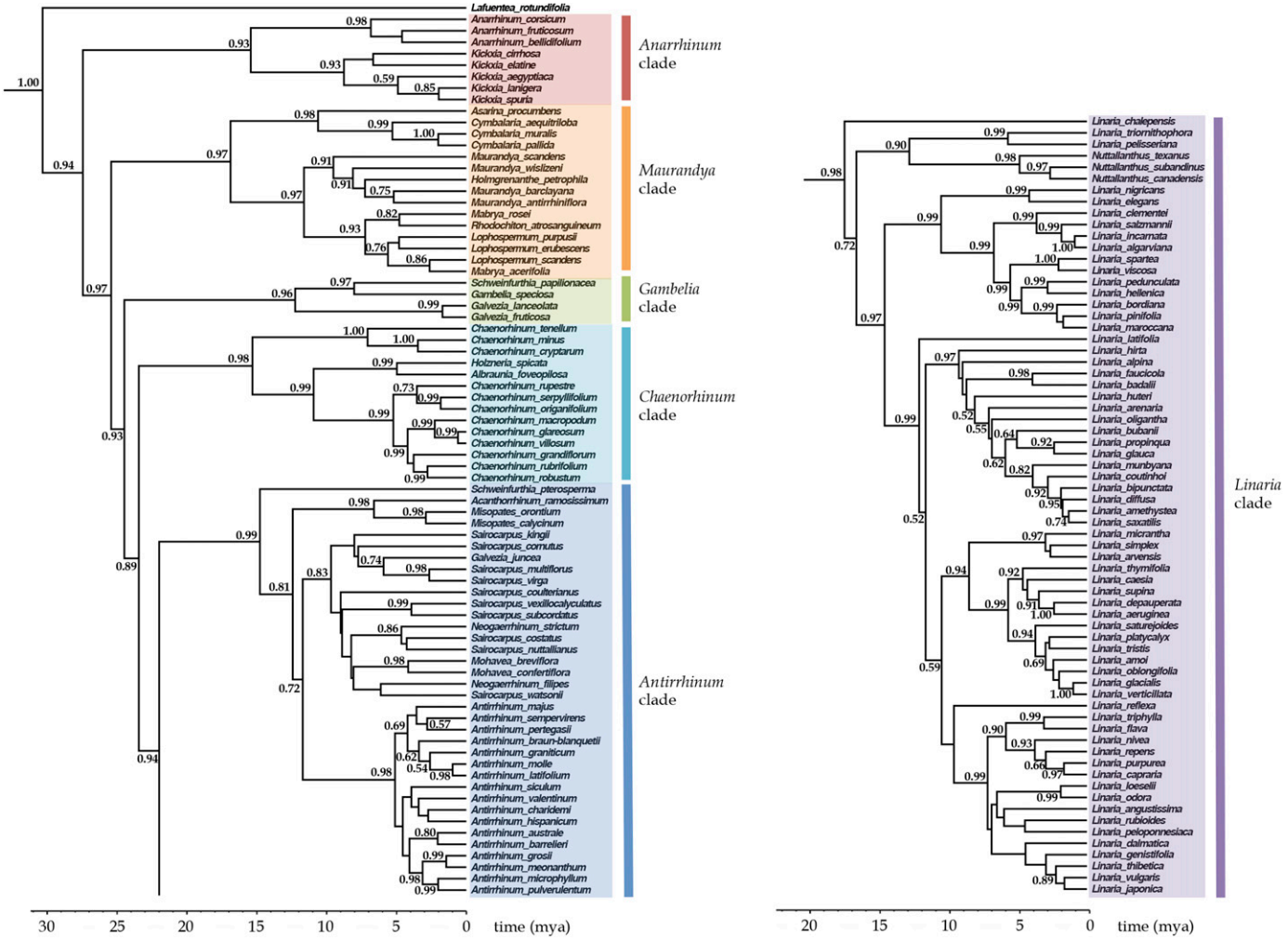


FIGURE 2 Maximum clade credibility tree for Antirrhineae. All five DNA regions were used to reconstruct the tree. Node labels show posterior probability (PP) values from the Bayesian analysis; PP values smaller than 0.50 are not shown. Age scale is in million years (mya), and six clades are highlighted in different colors.

(BS = 33%), both of which form the sister group to the *Antirrhinum* clade (BS = 68%).

All the differences between the Bayesian and the ML analyses, and between the different data sets, were observed at the nodes with low support values. To use the most inclusive Antirrhineae phylogeny with time-calibrated nodes, all the following analyses were performed with the Bayesian MCC tree constructed with the 5All sequence data set.

Estimated from the time-calibrated BEAST analysis, the node age for Antirrhineae was found to be 26.49 myr in the late Oligocene. The genus *Lafuentea*, which was previously shown to be a sister taxon to Antirrhineae (Albach et al., 2005), was represented with one species (*L. rotundifolia* Lag.) in our phylogeny, and was found to be closely related to the tribe (1.00 PP).

Geographic distribution and dispersal events—Having an Old World distribution was more common than a New World distribution in Antirrhineae. Out of the 146 species, 114 have an Old World distribution (78.08%), whereas the New World distribution is observed in 32 species (21.92%). The majority of the studied species were distributed in the Mediterranean Basin. When the distribution data were classified into three groups, the distribution of the 146 species was as follows: 32 species in the Americas (21.92%), 105 species in the Mediterranean (71.92%), and 34 species in the non-Mediterranean Old World (23.29%). Of these species, 25 were found both in the Mediterranean and the non-Mediterranean Old World (17.12%).

Results from both S-DIVA and DEC analyses agreed with the Bayesian analysis, therefore, only the Bayesian results are shown. Ancestral state reconstructions with binary state distribution data indicate an Old World origin (PP > 0.99), and the multiple state distribution data indicate a Mediterranean origin for Antirrhineae (PP > 0.99). According to the dispersal-vicariance analysis, the tribe has experienced five long-distance dispersal events between the Old World and the New World (Fig. 3; See Discussion section for vicariance vs. long-distance dispersal).

According to our divergence times estimation, all five long-distance dispersal events occurred in Antirrhineae during the Miocene: the first one occurred approximately 23.01 mya when the *Gambelia* clade was separated from the *Chaenorhinum, Antirrhinum*, and *Linaria* clades. The second long-distance dispersal was the separation of *Asarina* and *Cymbalaria* from the rest of the *Maurandya*

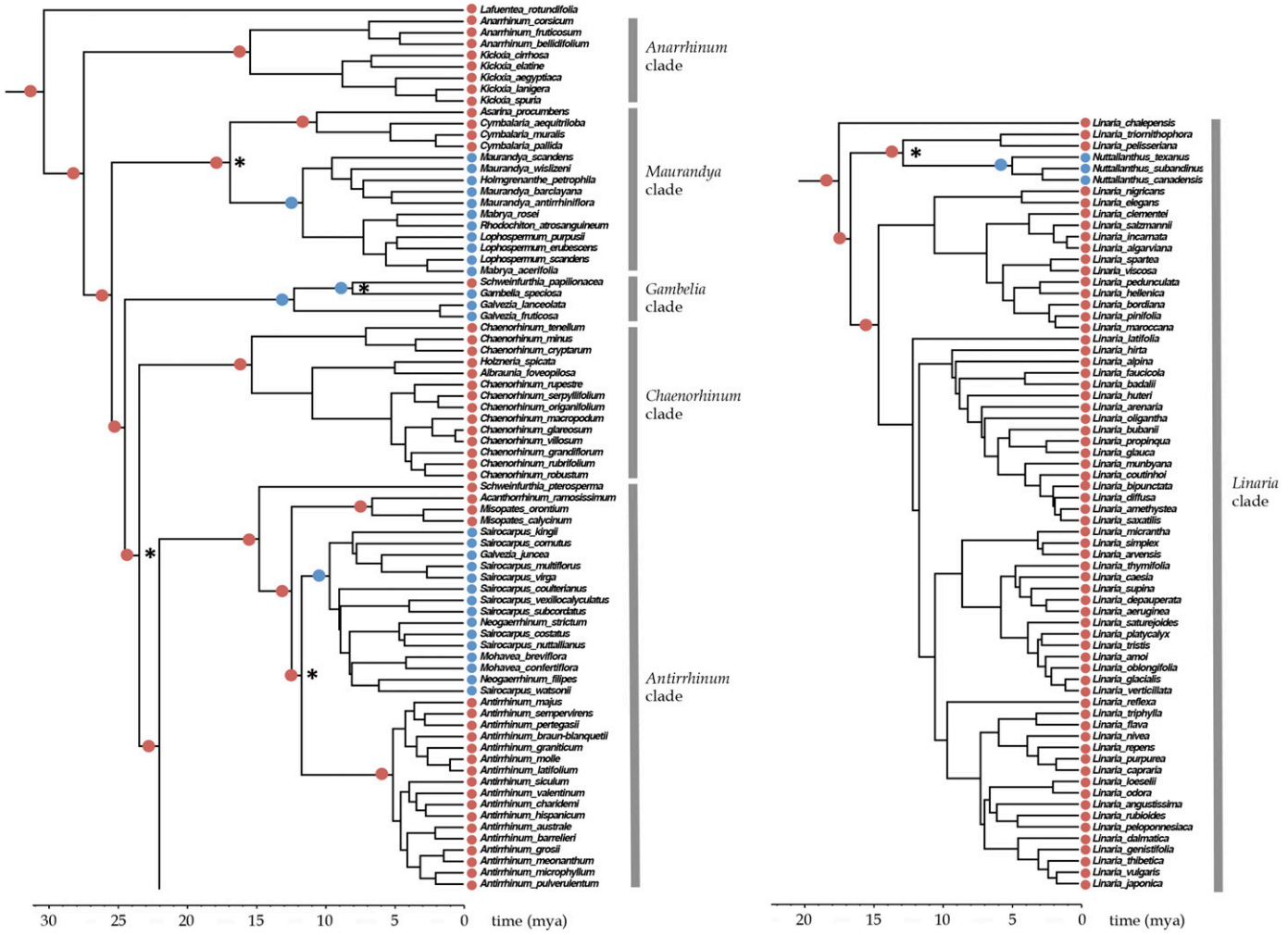


FIGURE 3 Bayesian Dispersal-Vicariance analysis results for the binary-state geographic distribution data. Blue: New World, Red: Old World. An asterisk (*) indicates nodes that experienced long-distance dispersal. The age scale is in million years (mya).

clade around 16.25 mya. The third long-distance dispersal occurred around 11.21 mya when *Nuttallanthus* separated from *Linaria*. The fourth long-distance dispersal was the separation of the *Maurandya-Neogaerrhinum-Sairocarpus* lineage from *Antirrhinum* around 9.99 mya. All four events represent long-distance dispersals from the Old World to the New World. The most recent long-distance dispersal is the separation of *Schweinfurthia papilionacea* from the rest of the *Gambelia* clade around 7.55 mya. This is the only long-distance dispersal from the New World to the Old World in the tribe.

Changes in ploidy levels—Both ML and Bayesian approaches inferred the ancestral haploid chromosome number for the tribe as 6 (Fig. 4). Overall, the tribe has experienced 11 gain events, 5 partial duplication events, and 7 duplication events (Fig. 4). The *Anarrhinum* clade experienced a partial duplication event, which increased the haploid chromosome number of the clade from 6 to 9, except for *Kickxia elatine* (L.) Dumort., which has $n = 18$ as a result of a recent duplication event. Most of the *Maurandya* clade has a haploid chromosome number of 12 as a result of an early duplication event. *Cymbalaria muralis* G.Gaertn., B.Mey. & Scherb. and *C. pallida* Wettst. have $n = 7$ after gain events, *C. aequitriloba* (Viv.) A.Chev. has $n = 28$ after two subsequent duplication events, and *Asarina procumbens* Mill. has $n = 9$ after a partial duplication event. The *Gambelia* clade experienced one partial duplication and one gain event at early stages, and the base chromosome number for the clade increased from $n = 6$ to $n = 10$. *Schweinfurthia papilionacea* went through another gain event and became $n = 11$, and *G. speciosa* became $n = 15$ after a partial duplication event. The haploid chromosome number of the *Chaenorhinum* clade is 7 as a result of a gain event. In the *Antirrhinum* clade, *Antirrhinum* has a haploid chromosome number of 8 as a result of 2 subsequent gain events. The *Sairocarpus-Mohavea* lineage experienced an additional partial duplication and a gain event and became $n = 15$, except for *Sairocarpus multiflorus* D.A.Sutton, *S. virga* (A.Gray) D.A.Sutton, and *S. cornutus* (Benth.) D.A.Sutton, which are all $n = 16$ as a result of an additional gain event. The *Linaria* clade remained $n = 6$ except for *Linaria chalepensis* (L.) Mill. and *L. pelisseriana* (L.) Mill., which have $n = 12$ as a result of duplication events.

DISCUSSION

The phylogeny of Antirrhineae—With the use of multiple markers of nuclear and plastid origin, and 146 species from 24 genera, our

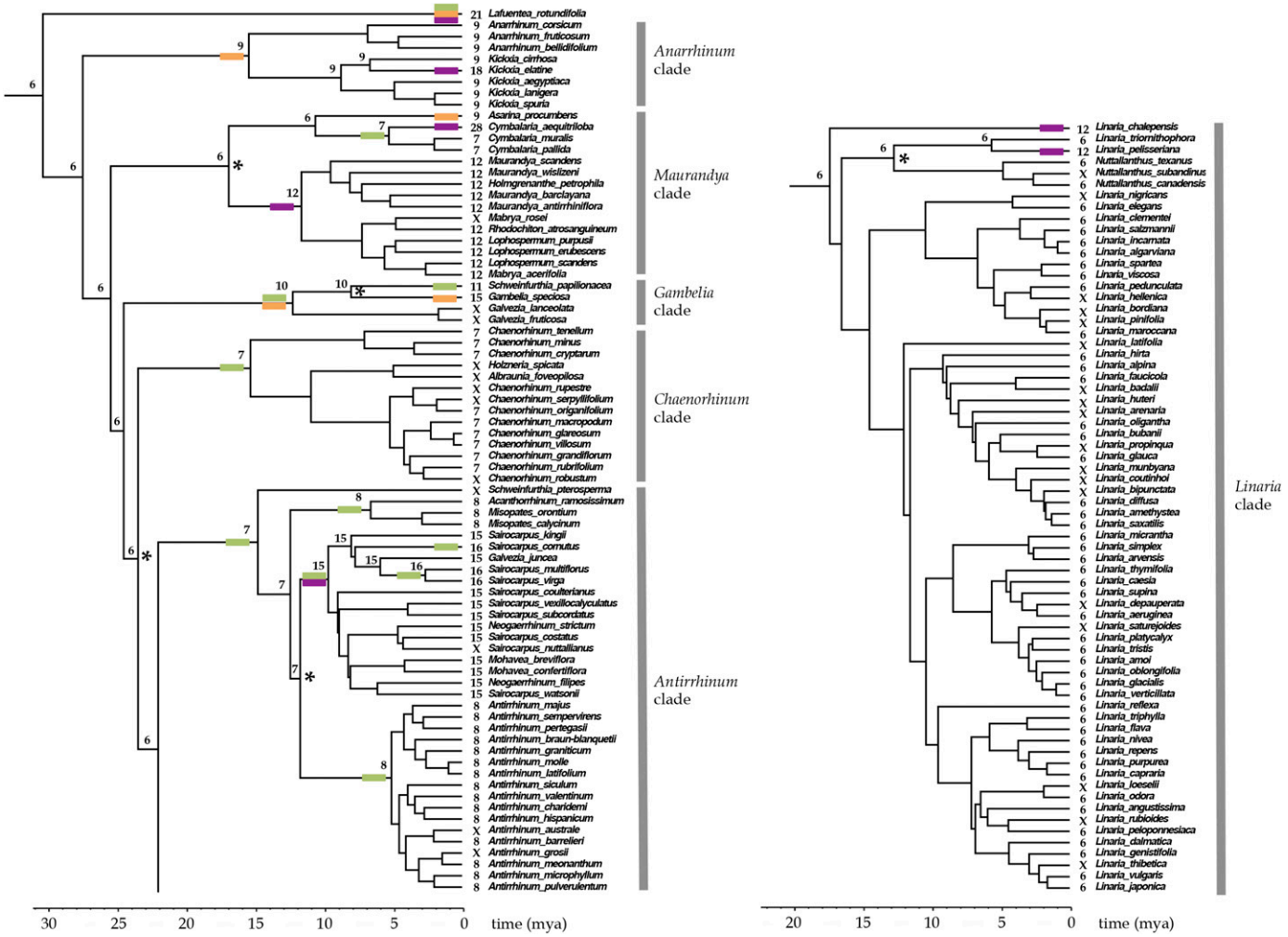


FIGURE 4 Chromosome number analysis results. Haploid chromosome numbers for the extant species are shown next to the species name. The numbers at the nodes state the inferred ancestral haploid chromosome number with the highest posterior probability and bootstrap values (support values are not shown). Purple lines indicate duplication events, orange lines indicate partial duplication events, and green lines indicate gain events. Unknown chromosome numbers are labeled as ×. An asterisk (*) indicates nodes that experienced long-distance dispersal. The age scale is in million years (mya).

sampling provided a detailed phylogenetic study for Antirrhineae, and resolved the monophyly of the tribe with six well-supported clades. Earlier studies of Antirrhineae phylogeny used either a single nuclear marker (Vargas et al., 2004; Fernández-Mazuecos et al., 2013), or a single plastid marker combined with morphological data (Ghebrehiwet et al., 2000). The most recent Antirrhineae phylogeny was built using one nuclear region (*ITS*) and one plastid region (*ndhF*), and included more taxa than previous studies (Guzman et al., 2015). It was previously shown that increasing the number of DNA regions improves the power of phylogenetic studies more than increasing the number of sampled taxa (Rokas and Carroll, 2005). Our study further increased the number of sequences sampled from one nuclear region (*ITS*) and one plastid region (*ndhF*) region, to one nuclear region (*ITS*) and four plastid regions (*ndhF, rbcL, rps16, trnL-F).*

Despite the use of multiple DNA regions and extensive taxon sampling, some branches still had low support values. When compared to likelihood methods, Bayesian analyses perform better phylogenetic reconstructions with missing data (Simmons, 2012). Another difference between the Bayesian approach and the likelihood methods is their support values. Posterior probability values tend to be higher than bootstrap values, but the correlation between them can be weak and variable (Douady et al., 2003). Posterior probability values tend to provide high support to clades with low phylogenetic signal, even though they also tend to give more type I errors (Erixon et al., 2003). For these reasons, we preferred using phylogenetic trees generated by Bayesian methods for our downstream analyses.

Our Antirrhineae phylogeny shares overall similarities with the previously published phylogenies that used alternative methods and had different sample sets (Ghebrehiwet et al., 2000; Vargas et al., 2004; Fernández-Mazuecos et al., 2013; Guzman et al., 2015). When we compare our phylogeny to the most recently published one, there are three main differences. (1) In terms of the relationships among the clades, the *Linaria* and *Antirrhinum* clades are sister taxa, and the *Chaenorhinum* clade is the sister group for these two clades in our phylogenetic tree. However, the previous study placed the *Linaria* clade as the sister group for the *Antirrhinum* and

the *Chaenorhinum* sister clades (Guzman et al., 2015). (2) Within the *Antirrhinum* clade, the organization of the genera is different between the two studies. The current phylogeny places *Antirrhinum* and the *Maurandya-Neogaerrhinum-Sairocarpus* lineage as sister taxa, and the *Acanthorrhinum-Misopates* lineage as their sister group, whereas in the previous phylogeny, *Antirrhinum* is the outgroup for the *Maurandya-Neogaerrhinum-Sairocarpus* and *Acanthorrhinum-Misopates* lineages (Guzman et al., 2015). (3) The placement of the *Galvezia*, *Gambelia*, and *Schweinfurthia* species differs between the two studies. The previous study places *Gambelia* species (*G. speciosa* and *G. juncea* (Benth.) D.A.Sutton) into the *Antirrhinum* clade, and the *Galvezia* clade (called *Gambelia* clade in this study) includes three *Galvezia* (*G. fruticosa* J.F.Gmel., *G. lanceolata* Pennell, *G. balli* Munz) and five *Schweinfurthia* (*S. pedicellata* Benth. & Hook.f., *S. spinosa* A.G.Mill., M.Short & D.A.Sutton, *S. latifolia* Baker, *S. papilionacea* Boiss., *S. imbricata* A.G.Mill., M.Short & D.A.Sutton*)* species. Our study did not include *G. balli, S. pedicellata, S. spinosa, S. latifolia,* and *S. imbricata* because of the species' unresolved status according to the Plant List (http://www.theplantlist.org). *Gambelia juncea* is a synonym for *Galvezia juncea*, and similar to the previous study, this species is placed into the *Antirrhinum* clade in the current phylogeny. The other *Gambelia* species we sampled was *G. speciosa*, and it is included in the *Gambelia* clade in our study. *Schweinfurthia pterosperma* was not sampled in the previous study, and our study places this species into the *Antirrhinum* clade, as an outgroup to the rest of the clade.

The New World genus *Nuttallanthus* was separated from the Old World genus *Linaria* and recognized as a different genus based on its distinctive floral and seed morphology (Sutton, 1988). *Nuttallanthus* shares the same chromosome number with *Linaria* ($n = 6$), and it is nested within this large genus *Linaria* in the current phylogeny. Accepting *Nuttallanthus* as a distinct genus would make *Linaria* a paraphyletic group. Therefore, *Nuttallanthus* should be considered as a subgroup of *Linaria* as suggested in an earlier study (Fernández-Mazuecos et al., 2013).

The *Gambelia/Galvezia* clade circumscription has been different in all the Antirrhineae phylogenies reconstructed so far (Ghebrehiwet et al., 2000; Vargas et al., 2004; Fernández-Mazuecos et al., 2013; Guzman et al., 2015). In the two most recent phylogenies, *Gambelia* was placed under the "*Antirrhinum* group", and *Galvezia* and *Schweinfurthia* formed their own monophyletic group (Fernández-Mazuecos et al., 2013; Guzman et al., 2015). Another phylogeny treated *Gambelia* and *Schweinfurthia* as a monophyletic group (no *Gambelia* species were sampled in the study), and put *Galvezia* under the "*Antirrhinum* group" (Vargas et al., 2004). Lastly, an earlier phylogeny for the tribe had *Gambelia* and *Galvezia* grouped together and *Schweinfurthia* clustered with the "*Antirrhinum* group" (Ghebrehiwet et al., 2000). Our phylogeny places the three genera in a monophyletic group, except for two species (*Galvezia juncea*, and *Schweinfurthia pterosperma*), which are included in the *Antirrhinum* clade. When we consider the geographic distribution and the chromosome numbers, the most parsimonious placement of these three genera would be separating the Old World genus *Schweinfurthia,* which is $n = 11$, from the New World genera *Gambelia* and *Galvezia*, both of which are $n = 15$. However, we do not have data to support such a phylogenetic placement. Future investigations with more markers would be helpful in elucidating the relationship among these three genera.

The geographic origin of Antirrhineae and long-distance dispersals from the Old World to New World—Found both in the Old World and the New World, Antirrhineae has a worldwide distribution, mostly concentrated in the Mediterranean region. Our ancestral reconstruction analysis showed a Mediterranean origin for the tribe, confirming the previous studies that show Antirrhineae originated in the Old World (Vargas et al., 2004; Vargas et al., 2014). Our time calibration data, which were based on the most recent study that estimated the divergence times across the angiosperms incorporating fossil data (Bell et al., 2010), were very similar to what Vargas et al. (2014) published, placing the Old World to New World dispersal of Antirrhineae in the Miocene.

According to our Bayesian binary state analysis, Antirrhineae experienced five "vicariance" events. It is problematic that the analysis is unable to differentiate between vicariance and long-distance dispersal events. Vicariance is defined as a splitting of a species' geographic range by a barrier that prevents gene flow between the split parts. By the time Antirrhineae was formed (~30 mya), both North America and Eurasia, and South America and Africa were separated (Raven and Axelrod, 1974). This makes the vicariance scenario less plausible, leaving long-distance dispersal as the more likely explanation for the range expansion of the tribe. Similar findings have been previously published for Antirrhineae (Fernández-Mazuecos and Vargas, 2011; Vargas et al., 2014). Transoceanic dispersals have been suggested to be the main factor shaping the distribution of many plant taxa today (McGlone et al., 2001; Rønsted et al., 2002; Givnish et al., 2004). It is still not clear how plants achieved long-distance dispersals across the ocean. Birds carrying seeds that are stuck to their bodies were suggested as one method (Rønsted et al., 2002), and wind anomalies resulting in storms due to the positive feedback loop between the oceanic winds and ocean temperatures was another explanation (Givnish et al., 2004).

A recent study found a relationship between polyploidy and long-distance dispersal, suggesting that polyploids have better success at long-distance dispersals than diploids (Linder and Barker, 2014). This supports the idea that polyploid lineages may have higher survival success in new environments (Stebbins, 1984; Brochmann et al., 2004). While there have been several reports of high numbers of polyploid taxa at high latitudes and altitudes (Vamosi and McEwen, 2013; Brochmann et al., 2004), few studies have considered new environments in the context of long-distance dispersal. In our study, except for the *Linaria-Nuttallanthus* differentiation, during which the haploid chromosome number remained as $n = 6$, all the long-distance dispersals occurred in conjunction with some changes in chromosome numbers. Separation of the New World group (*Holmgrenanthe, Lophospermum, Mabrya, Maurandya, Rhodochiton*) in the *Maurandya* clade from the Old World group (*Asarina, Cymbalaria*) is followed by a duplication event, increasing the haploid chromosome number from 6 to 12. Similarly, the separation of the New World genera (*Sairocarpus, Mohavea, Neogaerrhinum*) in the *Antirrhinum* clade from the rest of the Old World genera (*Acanthorrhinum, Antirrhinum, Misopates*) is followed by a change in haploid chromosome numbers from 7 to 15 via duplication and gain events. The *Gambelia* clade experienced two long-distance dispersal events, both of which were associated with changes in the chromosome numbers. The first long-distance dispersal was from the Old World to the New World, and it was coupled with partial duplication and gain events, increasing the haploid chromosome number from 6 to 10. The second long-distance dispersal was to the opposite direction:

from the New World to the Old World, and it was coupled with a gain event.

The ability of polyploids to colonize new areas has previously been suggested to be due to their higher allelic diversity, which gives them flexibility to adapt to new environments (Stebbins, 1940; Soltis et al., 1993). In the current study, when dispersal and chromosome number changes are observed concurrently, the common pattern is that the dispersal is followed by a change in chromosome number. This pattern contradicts the idea that polyploids are more successful in colonizing new areas, because the timeline of events suggests that dispersal is followed by polyploidization, not the other way. Polyploid plants are observed more commonly in areas that have not been previously occupied, whereas diploid plants more often occupy stable habitats that were established a long time ago (Stebbins, 1940). By revisiting this idea, it can be suggested that polyploids are observed more often in unoccupied environments, either because the polyploid taxa have a better chance of survival there, or alternatively, the polyploidization occurs more often in these areas than anywhere else.

In conclusion, the tools are now available to examine the contribution of long-distance dispersal and genome evolution to macroevolutionary patterns within a large range of clades. Here, with 24 recognized genera represented by 146 species and five DNA regions, we reconstructed an extensive Antirrhineae phylogeny. While chromosome patterns have often been previously explained by the evolution of polyploid taxa during the advances and retreats of ice sheets during Pleistocene (Stebbins, 1984), we find that long-distance dispersal and genome duplication reconstructions help circumscribe the tribes. Further analyses are needed to address the effects of these dispersal patterns on the diversification of Antirrhineae.

ACKNOWLEDGEMENTS

The authors thank J. Theriault for helping them with the data matrix construction, and the two anonymous reviewers for their valuable suggestions and comments. This work was supported by an NSERC Discovery Grant to JCV.

LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–723.

Albach, D. C., H. M. Meudt, and B. Oxelman. 2005. Piecing together the "New" Plantaginaceae. *American Journal of Botany* 92: 297–315.

Baldwin, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molecular Phylogenetics and Evolution* 1: 3–16.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. *Systematic Biology* 51: 625–637.

Bell, C. D., D. E. Soltis, and P. S. Soltis. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* 97: 1296–1303.

Brochmann, C., A. K. Brysting, I. G. Alsos, L. Borgen, H. H. Grundt, A. C. Scheen, and R. Elven. 2004. Polyploidy in arctic plants. *Biological Journal of the Linnean Society. Linnean Society of London* 82: 521–536.

Darlu, P., and G. Lecointre. 2002. When does the incongruence length difference test fail? *Molecular Biology and Evolution* 19: 432–437.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9: 772.

De Bodt, S., S. Maere, and Y. Van de Peer. 2005. Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution* 20: 591–597.

Dong, W., T. Cheng, C. Li, C. Xu, P. Long, C. Chen, and S. Zhou. 2014. Discriminating plants using the DNA barcode *rbcLb*: an appraisal based on a large data set. *Molecular Ecology Resources* 14: 336–343.

Douady, C. J., F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. P. Douzery. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution* 20: 248–254.

Doyle, J.J., and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

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

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* 52: 665–673.

Escudero, M., S. Martín-Bravo, I. Mayrose, M. Fernández-Mazuecos, O. Fiz-Palacios, A.L. Hipp, M. Pimentel, et al. 2014. Karyotypic Changes through dysploidy persist longer over evolutionary time than polyploid changes. *PLoS One* 9: e85266.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994. Testing significance of congruence. *Cladistics* 10: 315–319.

Fernández-Mazuecos, M., and P. Vargas. 2011. Historical isolation versus recent long-distance connections between Europe and Africa in bifid toadflaxes (*Linaria* sect. *Versicolores*). *PLoS One* 6: e22234.

Fernández-Mazuecos, M., J. L. Blanco-Pastor, and P. Vargas. 2013. A phylogeny of toadflaxes (*Linaria* Mill.) based on nuclear internal transcribed spacer sequences: systematic and evolutionary consequences. *International Journal of Plant Sciences* 174: 234–249.

Gernhard, T. 2008. The conditioned reconstructed process. *Journal of Theoretical Biology* 253: 769–778.

Ghebrehiwet, M., B. Bremer, and M. Thulin. 2000. Phylogeny of the tribe Antirrhineae (Scrophulariaceae) based on morphological and *ndhF* sequence data. *Plant Systematics and Evolution* 220: 223–239.

Givnish, T. J., K. C. Millam, T. M. Evans, J. C. Hall, J. C. Pires, P. E. Berry, and K. J. Sytsma. 2004. Ancient vicariance or recent long-distance dispersal? Inferences about phylogeny and South American–African disjunctions in Rapateaceae and Bromeliaceae based on *ndhF* sequence data. *International Journal of Plant Sciences* 165: S35–S54.

Glick, L., and I. Mayrose. 2014. ChromEvol: assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. *Molecular Biology and Evolution* 31: 1914–1922.

Guindon, S., and O. Gascuel. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* 52: 696–704.

Guzmán, B., J. M. Gómez, and P. Vargas. 2015. Bees and evolution of occluded corollas in snapdragons and relatives (Antirrhineae). *Perspectives in Plant Ecology, Evolution and Systematics* 17: 467–475.

Hassler, M. 2016. World Plants: synonymic checklists of the vascular plants of the world (version Nov 2015). *In* Roskov, Y., L. Abucay, T. Orrell, D. Nicolson, C. Flann, N. Bailly, P. Kirk, T. Bourgoin, DeWalt, R. E., W. Decock, and A. De Wever [eds.], Species 2000 & ITIS Catalogue of Life. Naturalis, Leiden, Netherlands.

Heled, J., and R. R. Bouckaert. 2013. Looking for trees in the forest: summary tree from posterior samples. *BMC Evolutionary Biology* 13: 221–231.

Linder, H., and N. Barker. 2014. Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113: 1175–1183.

Maddison, W. P., and D. R. Maddison. 2014. Mesquite: a modular system for evolutionary analysis. Version 3.0 http://mesquiteproject.org.

Mayrose, I., M. S. Barker, and S. P. Otto. 2010. Probabilistic models of chromosome number evolution and the inference of polyploidy. *Systematic Biology* 59: 132–144.

Mayrose, I., S. H. Zhan, C. J. Rothfels, K. Magnuson-Ford, M. S. Barker, L. H. Rieseberg, and S. P. Otto. 2011. Recently formed polyploid plants diversify at lower rates. *Science* 333: 1257.

McGlone, M. S., R. P. Duncan, and P. B. Heenan. 2001. Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand. *Journal of Biogeography* 28: 199–216.

Oxelman, B., M. Liden, and D. Berglund. 1997. Chloroplast *rpsl6* intron phylogeny of the tribe Sileneae (Caryophyllaceae). *Plant Systematics and Evolution* 206: 393–410.

Peters, R. S., B. Meyer, L. Krogmann, J. Borner, K. Meusemann, K. Schütte, O. Niehaus, and B. Misof. 2011. The taming of an impossible child: a standardized all-in approach to the phylogeny of Hymenoptera using public database sequences. *BMC Biology* 9: 55–68.

Rambaut, A., 2009. FigTree v1. 3.1: Tree figure drawing tool.

Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2014. Tracer v1.6. Available from: http://beast.bio.ed.ac.uk/Tracer.

Raven, P. H., and D. I. Axelrod. 1974. Angiosperm biogeography and past continental movements. *Annals of the Missouri Botanical Garden* 61: 539–673.

Ree, R. H., and S. A. Smith. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14.

Rice, A., L. Glick, S. Abadi, M. Einhorn, N. M. Kopelman, A. Salman-Minkov, J. Mayzel, et al. 2015. The chromosome counts database (CCDB)—a community resource of plant chromosome numbers. *The New Phytologist* 206: 19–26.

Rokas, A., and S. B. Carroll. 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. *Molecular Biology and Evolution* 22: 1337–1344.

Ronquist, F. 1997. Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46: 195–203.

Ronquist, F. 2004. Bayesian inference of character evolution. *Trends in Ecology & Evolution* 19: 475–481.

Rønsted, N., M. W. Chase, D. C. Albach, and M. A. Bello. 2002. Phylogenetic relationships within *Plantago* (Plantaginaceae): Evidence from nuclear ribosomal *ITS* and plastid *trnL-F* sequence data. *Botanical Journal of the Linnean Society* 139: 323–338.

Rothmaler, W. 1943. Zur Gliederung der Antirrhineae. *Feddes Repertorium Specierum Novarum Regni Vegetabilis.* 52: 16–39.

Rothmaler, W. 1956. Taxonomische Monographie der Gattung *Antirrhinum. Feddes Repertorium Specierum Novarum Regni Vegetabilis.* 136: 1–124.

Schwarz, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464.

Simmons, M. P. 2012. Misleading results of likelihood-based phylogenetic analyses in the presence of missing data. *Cladistics* 28: 208–222.

Soltis, D. E., P. S. Soltis, and L. H. Rieseberg. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.

Stamatakis, A. 2014. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)* 30: 1312–1313.

Stebbins, G. L. 1940. The significance of polyploidy in plant evolution. *American Naturalist* 74: 54–66.

Stebbins, G. L. 1971. Chromosomal evolution in higher plants. Edward Arnold, London London, UK.

Stebbins, G. L. 1984. Polyploidy and the distribution of the arctic alpine flora: new evidence and a new approach. *Botanica Helvetica* 94: 1–13.

Sutton, D. A. 1988. A revision of the tribe Antirrhineae. Oxford University Press, London, UK.

Swofford, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer. Sunderland, MA.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* 17: 57–86 (American Mathematical Society).

The Plant List. 2013. Web: http://www.theplantlist.org.

Vamosi, J. C., and T. A. Dickinson. 2006. Polyploidy and diversification: A phylogenetic investigation in Rosaceae. *International Journal of Plant Sciences* 167: 349–358.

Vamosi, J. C., and J. R. McEwen. 2013. Origin, elevation and evolutionary success of hybrids and polyploids in British Columbia, Canada. *Botany* 91: 182–188.

Vargas, P., J. A. Rosselló, R. Oyama, and J. Güemes. 2004. Molecular evidence for naturalness of genera in the tribe Antirrhineae (Scrophulariaceae) and three independent evolutionary lineages from the New World and the Old. *Plant Systematics and Evolution* 249: 151–172.

Vargas, P., L. M. Valente, J. L. Blanco-Pastor, I. Liberal, B. Guzman, E. Cano, A. Forrest, and M. Fernández-Mazuecos. 2014. Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: snapdragons and the Madrean–Tethyan flora. *Journal of Biogeography* 41: 932–943.

Wheeler, T. J., and J. D. Kececioglu. 2007. Multiple alignments by aligning alignments. *Bioinformatics (Oxford, England)* 23: i559–i568.

Wood, T. E., N. Takebayashi, M. S. Barker, I. Mayrose, P. B. Greenspoon, and L. H. Reiseberg. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences, USA* 106: 13875–13879.

Yu, Y., A. J. Harris, C. Blair, and X. J. He. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution* 87: 46–49.

APPENDIX 1 **Details about the dried plant materials collected from the University of Calgary Herbarium. The specimens were used to generate de novo sequences (see Materials and Methods for details).**

Species	GenBank accession number	Voucher ID	Collected from	Collection Date	Collector
Chaenorhinum minus	*ITS*: KX061032 *rbcL*: KX037372 *rps16*: KX037392 *trnL-F*: KX061013	55652	AB, Canada	2004	I.D. MacDonald & K. Hull
Kickxia lanigera	*ITS:* KX061033 *rbcL*: KX037373 *rps16:* KX037393 *trnL-F*: KX061014	44007	Costa del Sol, Spain	1983	B.M. Hallworth
Kickxia spuria	*rps16*: KX037397 *trnL-F*: KX061018	38009	Costa del Sol, Spain	1980	B.M. Hallworth
Linaria dalmatica	*ITS:* KX061035 *rbcL*: KX037377 *rps16:* KX037398 *trnL-F*: KX061019	51111	AB, Canada	1995	K. Sharpe
Linaria repens	*rps16:* KX037411	36897	NL, Canada	1978	J.K. Morton & J. M. Venn
Linaria vulgaris	*ITS:* KX061036 *rbcL*: KX037378 *rps16:* KX037399 *trnL-F*: KX061020	49832	AB, Canada	1988	S. Rai