Phylogeny, biogeography, and rates of diversification of New World *Astragalus* (Leguminosae) with an emphasis on South American radiations[1]

Rosa A. Scherson,[2,4] Rodrigo Vidal,[3] and Michael J. Sanderson[2,5]

[2]Section of Evolution and Ecology, University of California, Davis, California 95616 USA; and [3]Molecular Ecology and Evolutionary Studies Laboratory, Department of Chemistry and Biology, Universidad de Santiago, Chile

This study uses phylogenetic relationships of New World representatives of the species-rich genus *Astragalus* (Leguminosae; Papilionoideae) to follow up on recent evidence pointing to rapid and recent plant diversification patterns in the Andes. Bayesian and maximum likelihood phylogenetic analyses were done using nuclear rDNA ITS and chloroplast spacers *trnD-trnT* and *trnfM-trnS1*, either separately or in combination. The effect of using partitioned vs. nonpartitioned analyses in a Bayesian approach was evaluated. Highest resolution was obtained when the data were combined in partitioned or nonpartitioned Bayesian analyses. All phylogenies support two clades of South American species nested within the North American species, implying two separate invasions from North to South America. These two clades correspond to the original morphological classification of Johnston (1947 *Journal of the Arnold Arboretum* 28: 336–409). The mean ages of the South American clades were very recent but still significantly different (1.89 and 0.98 Ma). Upper and lower bounds on rates of diversification varied between 2.01 and 0.65 species/Ma for the older clade and 2.06 and 1.24 species/Ma for the younger clade. Even the lower bounds are still very high, reasserting Neo-*Astragalus* in the growing list of recent rapid radiations of plants, especially in areas with a high physiographic diversity, such as the Andes.

Key words: *Astragalus*; biogeography; chloroplast spacers; internal transcribed spacer; Leguminosae; phylogeny; rates of diversification.

The increasing availability of time-calibrated molecular phylogenies is uncovering surprisingly high rates of species diversification in many clades (Richardson et al., 2001; Allan et al., 2004; Klak et al., 2004; Near and Benard, 2004; Hughes and Eastwood, 2006). In legumes, the third largest family of angiosperms (Mabberley, 1997), examples include *Lotus* L., in which rapid diversification was suggested after colonization of the Canary Islands (Allan et al., 2004), *Inga* Mill. with a radiation of more than 300 species in neotropical wet forests, diversifying as recently as 2 million years ago (Ma) (Richardson et al., 2001), and the Andean tropical clade of *Lupinus* L., with island-like diversification rates of 2.49–3.72 net new species per million years (Hughes and Eastwood, 2006), much higher than rates inferred for true island radiations such as the Hawaiian silversword alliance (Baldwin and Sanderson, 1998).

The papilionoid legume taxon *Astragalus* L. is among the most species-rich plant genera with more than 2500 species inhabiting Mediterranean arid and semiarid environments (Lock and Simpson, 1991; Mabberley, 1997; Massoumi, 1998; ILDIS, 2002). Because of its large numbers of species it is very suitable for studying processes of species diversification at a large scale. Among the best-supported clades within *Astragalus* is Neo-

Astragalus, with some 500 New World aneuploid species nested within an older and more diverse euploid assemblage of other taxa (Wojciechowski et al., 1993, 1999). Johnston (1947) recognized nearly 90 species in South America, and in more recent works, between 108 and 110 species have been recognized (Gómez-Sosa, 1994, 2003; ILDIS, 2002), with distributions from Ecuador to the Strait of Magellan and two main centers of speciation, one in the high country of northwestern Argentina and adjacent Bolivia and the other in the Andes in Chile and Argentina south of Aconcagua (Johnston, 1947). They are also found in arid coastal regions (Johnston, 1947; Gómez-Sosa, 1994). No species of *Astragalus* is present in both North and South America.

Diversification rates in the genus *Astragalus* as a whole and among its outgroups in the Astragalean clade (Fig. 1) have been addressed previously (Liston, 1989; Sanderson and Wojciechowski, 1996; Wojciechowski et al., 1999), but there is still no consensus on whether *Astragalus* or any of its subclades have some intrinsic characteristic that has facilitated diversification (e.g., predisposition to the evolution of intrinsic barriers to reproduction or unusually high amounts of genetic diversity). Analyses have been hampered by at least two factors. First, molecular resolution within the clade is poor due to a lack of phylogenetically informative characters. Second, there is a lack of proper calibration points to obtain age estimates, due to poor representation in the fossil record. Recently, a comprehensive study of evolutionary rates within the Leguminosae has allowed accurate estimation of the times of diversification of all its major clades (Lavin et al., 2005). Subsequent analyses provided age estimates for the Astragalean clade (12.4 ± 1.45 Ma) and Neo-*Astragalus* (4.4 ± 0.78 Ma: Wojciechowski, 2005).

Previous phylogenetic work placed sampled South American species of Neo-*Astragalus* into two separate clades (Wojciechowski et al., 1999), which correspond to two taxa in the only subgeneric classification of the South American
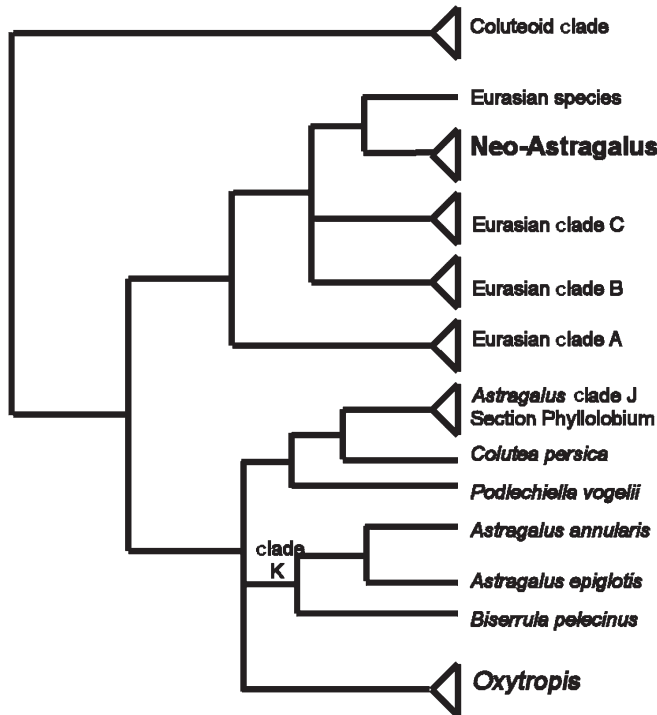
Fig. 1. Schematic representation of the Astragalean clade. Relative position of the Coluteoid clade from Wojciechowski et al. (2004); major clades within *Astragalus* and *Astragalus* s.s. and relative position of *Oxytropis* from Osaloo et al. (2005); relative position of the Neo-Astragalus clade from Wojciechowski et al. (1999).

species (Johnston, 1947): section *Phaca* with 15 species that are distinguished by separated, herbaceous stipules and section *Euastragalus* with more or less united, membranous stipules, which includes the remaining species in South America. One of the clades in the phylogeny of Wojciechowski et al. (1999) included species from North America, signaling a pattern of amphitropical disjunction. However, lack of phylogenetic resolution within the group has not permitted a proper test of the nature of that disjunction: whether it occurs at the level of sister clades or if it is a disjunction of sister species within a mixed North-South American clade.

Phylogenetic analyses in Neo-Astragalus have been especially difficult due to low levels of molecular polymorphism. A large sample of nuclear genes screened so far have proven problematic for phylogenetic reconstruction because of duplication events in some species and reduced levels of sequence divergence relative to the divergence in Old World *Astragalus* (Scherson et al., 2005).

Here we use two chloroplast spacers not used previously in *Astragalus*. With an improved taxon sample, better phylogeny and recently reported secondary age calibrations for Neo-Astragalus, we set out to test the hypothesis that the South American clades of Neo-Astragalus have high rates of species diversification, as seen in other Andean plant radiations. In addition, we test the use of partitioned vs. nonpartitioned approaches to phylogenetic analysis in a Bayesian framework, to investigate whether the use of different models in a combined data set helps to increase resolution. Biogeographical hypotheses on the nature of the disjunct distribution of Neo-Astragalus are also discussed.

MATERIALS AND METHODS

Taxon sampling—Forty-eight species of *Astragalus* were sampled, including 17 South American species covering the complete range of its distribution, from Ecuador to Patagonia and from the high Andes to the coastal dunes; and 27 North American members of the aneuploid Neo-Astragalus clade, including six varieties of *A. lentiginosus* Douglas ex Hook. and two of *A. allochrous* A. Gray. Two euploid New World species, *A. americanus* (Hook.) M. E. Jones, and *A. canadensis* L., the euploid species *A. adsurgens* Pall., distributed in North America and Asia, in addition to *Oxytropis sericea* Nutt.ex Torr. were used as outgroups. Species names, vouchers, and GenBank accessions are detailed in Appendix 1.

Molecular markers and DNA amplification—Two chloropast spacers (*trnD-trnT* and *trnfM-trnS1*) located in the large single copy (LSC) region of the chloroplast were selected after five chloroplast regions were screened according to the criteria of Shaw et al. (2005) for interspecific variation. Spacers were amplified by PCR using the conditions and primers described in Shaw et al. (2005). For the spacer *trnfM-trnS1*, a touch-down PCR profile was used to eliminate nonspecific product. The profile was set as follows: 94°C for 3 min; 9 cycles of 94°C, 70°C, 67°C, and 72°C, each for 45 s; 5 cycles of 94°C, 67°C, 65°C, and 72°C, each for 45 s; 15 cycles of 94°C, 65°C, and 72°C each for 45 s; and a final extension of 72°C for 3 min.

The 5.8S subunit and flanking internal transcribed spacer (ITS) of ribosomal DNA, previously shown to be of utility in *Astragalus* (Wojciechowski et al., 1999) was also used. Sequences for 26 species were obtained from previous work by Wojciechowski et al. (1993, 1999). The ITS sequence for the remaining 22 species was obtained by amplification using the primers ITS4 and ITS5 (White et al., 1990) according to Wojciechowski et al (1993). PCR products were purified using EXOSAP (USB Corp., Cleveland, Ohio, USA) and sequenced at the University of California-Davis, Division of Biological Sciences sequencing facility. The resulting chromatograms were analyzed using the program Sequencher version 4.1 (Gene Codes Corp., Ann Arbor, Michigan).

Phylogenetic analyses—DNA sequences were aligned manually using the programs Se-Al version 2.0a1 (Rambaut, 1996) and Seaview (Galtier et al., 1996). Three data matrices were obtained: (1) ITS alone ("ITS"), (2) the two chloroplast spacers together ("chloroplast") and (3) ITS plus chloroplast ("combined").

Models of sequence evolution were selected using the program MrModeltest version 2.2 (Nylander, 2002) based on the Akaike information criterion (AIC) (Posada and Buckley, 2004). On the basis of this analysis, data sets were analyzed using the GTR+I+G model for the chloroplast dataset and the SYM+G model for ITS. The combined data set was analyzed as one partition using the GTR+I+G model or in two partitions with both models combined. AIC scores for the five best models for each partition are shown in Appendix S1 (see Supplemental Data with the online version of this article).

The program MrBayes version3.1 (Ronquist and Huelsenbeck, 2003) was used for the Bayesian phylogenetic analyses. For the partitioned analysis, the substitution and branch length estimates were allowed to vary independently in each partition. Posteriors on the model parameters were estimated from the data, using the default priors. The separate data set analyses were done using 10 million generations and the combined data set analyses were done with 20 million generations, sampling the Markov chain every 100 generations, and using four independent chains running simultaneously. The number of generations was adjusted to ensure stationarity of the Markov chain based on the following indicators: (1) a stable value of the log likelihood of the cold chain in two separate runs (Appendix S2, see Supplemental Data with the online version of this article); (2) a value approaching zero for the standard deviation between runs (0.006 for both runs); and (3) a value approaching 1.0 for the potential scale reduction factor (PSRF) for each parameter in the model. The trees obtained were imported into the program PAUP* (Swofford, 2001), and a majority rule consensus tree was computed after discarding the first 25% of the trees (burnin), which were saved prior to MCMC convergence. Support for clades given by posterior probabilities was thus represented by the majority rule percentage.

A maximum likelihood bootstrap analysis with two hundred replicates was done using PAUP* (Swofford, 2001) for the combined data set using the GTR+I+G model. The substitution parameters were estimated over a neighbor-joining tree. Each replicate was run using a heuristic search with TBR branch swapping and limiting the search time to 12 h per replicate.

Evolutionary rates and dating analyses—Analyses of ages and rates of diversification were done in the context of a credible set of nonpartitioned Bayesian phylogenies. Some 145 000 trees with their estimated branch lengths were

imported into PAUP* (Swofford, 2001), after removing the first 25% of them to account for burnin. Credible intervals of posterior distributions were obtained using 1000 of those trees, randomly chosen. The trees were filtered according to a constraint tree that contained the clades of interest, namely Neo-Astragalus and the South American clades, without constraining their internal topologies. All trees sampled contained these three clades, which are highly supported (see Results).

Evolutionary rates analyses—The set of 1000 trees was subsequently imported into the program r8s (Sanderson, 2003). The penalized likelihood method (Sanderson, 2002) was used to estimate nucleotide substitution rates and ages of the selected clades. This method allows evolutionary rates to vary smoothly between branches. A cross-validation test using 10 randomly selected trees was performed to find an optimum rate smoothing parameter, which generally falls between the single parameter model (molecular clock) and a parameter-rich model. The cross-validation was done using a range of 13 values in increments of 0.25. The truncated Newton (TN) optimization algorithm was used for this analysis (Sanderson, 2002). A credible interval for the divergence time of a node was constructed as ±2 SD in the distribution of ages seen across the whole posterior distribution of trees.

Due to the fact that the data effectively behaved in a clocklike manner (see Results for details), the Langley–Fitch molecular clock method was used to compute ages of the two nodes of interest for the set of trees using r8s (Sanderson, 2003). A *t* test was used to determine if the ages inferred for each node were significantly different from each other.

In addition, because r8s does not use the same substitution model as was used to infer the trees, an independent molecular clock analysis was carried out using PAUP* (Swofford, 2001). The same set of 1000 Bayesian trees was imported into PAUP* without branch lengths. Maximum likelihood was used to estimate the branch lengths enforcing a molecular clock, using the GTR+I+G model of sequence evolution as selected with MrModeltest version 2.2 (Nylander, 2002). The set of ultrametric trees generated in this manner was subsequently imported into r8s (Sanderson, 2003) and the 'calibrate' command was used to set the inferred ages of the South American clades on an absolute time scale.

Calibration for the divergence time analyses was based on a previously estimated age for the Neo-Astragalus crown node, of 4.4 ± 0.78 Ma (Wojciechowski, 2005), which has some uncertainty. Because we are also estimating ages with uncertainty, our final error estimates have two components, the first arising from the uncertainty in the calibration, the second from the uncertainty due to sequence evolution in our data. The overall uncertainty can be estimated using a standard formula for "propogation of errors" in a product of two variables (the squared overall fractional error is just the sum of the squared individual fractional errors) (Goodman, 1960).

Upper and lower bounds for ages and rates—Estimates of the ages and rates of diversification of clades are sensitive to incomplete sampling (Linder et al., 2005). If older species are left out, the crown group tends to appear too young, and the rates of diversification can be overestimated. This bias is especially important in large taxa (Magallón and Sanderson, 2001; Wojciechowski et al., 2005). Consequently, ages were calculated for both the crown node (most recent common ancestor of the sampled species of the clade of interest) and the stem node (most recent common ancestor of the clade of interest and its sister clade). In addition, analyses were done to account for possible errors in the net number of species per clade. Four rates of diversification were thus calculated, using crown vs. stem ages on the one hand, and for each of those, the number of species hypothesized to be the "real" number for each clade (N) vs. the actual number of sampled species (S).

Diversification rates were estimated assuming a Yule (pure birth) process:

$$r = [\ln(n) - \ln(n_0)]/t,$$

where r = rate of species diversification, n = number of species in the clade (N or S), t = age of crown or stem node, and n_0 = 1 for stem age and 2 for crown age.

RESULTS

Sequence analyses—Sequence identity for amplified ITS was confirmed by comparison to the sequences obtained from Wojciechowski et al. (1999). After introducing gaps and eliminating ambiguous characters, a 686-bp matrix of the 5.8S, ITS2 and 26S region was generated, with 88 variable characters, 46 of them phylogenetically informative.

For the chloroplast regions, sequence identity was confirmed by comparison to the corresponding loci in *Medicago truncatula* Gaertn. and *Glycine max* (L.) Merr. available in GenBank. For the *trnfM-trnS1* section, ~115 bp of the *trnF-trnG* spacer, the full *trnG* gene, ~500 bp of the *trnG-psbZ* spacer, the full *psbZ* gene, and ~50 bp of the *psbZ-trnS1* spacer were sequenced. Sequencing of the *trnD-trnT* spacer included the full *trnY* gene and about 670 bp of the spacer sequence. In both *trnT-trnD* and *trnfM-trnS1*, there was a section between 200 and 270 bp that could not be sequenced due to the presence of large polyA/T regions. The sequenced *trnfM-trnS1* region yielded a 1066 character matrix of unambiguous alignment, with 60 variable characters, 21 of them parsimony informative. The *trnD-trnT* was more variable, yielding a matrix of 1001 characters from which 110 characters were variable, and 28 parsimony informative.

Rates of substitutions—*Pseudo-clocklike behavior of the data*—A χ^2 test rejected a molecular clock for the combined data set ($\chi^2 = 119.425786$, df = 44), indicating that the absolute rate of substitutions differs across branches. A set of 10 randomly selected trees was then used to do a cross-validation analysis in r8s, which searches for the optimal smoothing value to be used in a penalized likelihood analysis. Ideally, the optimal value is selected by finding the minimum χ^2 error. Sometimes, however, the error is a constant or irregularly varying function of the smoothing parameter, indicating that the predictive power of the penalized likelihood model is no better than that obtained by assuming a clock. In this data set, the optimal smoothing values obtained from cross-validation differed for each tree; in some cases, there were several smoothing values showing the same minimal χ^2 error; and there was rarely much difference between predictive accuracy under optimally smooth penalized likelihood and a simple clock (Appendix S3, see Supplemental Data with the online version of this article). Even though the rate of substitution varies across branches, the rate-smoothed model has no better predictive accuracy than the Langley–Fitch (clock) algorithm. This lack of difference between the predictive capacity of the rate-smoothing vs. a molecular clock, despite varying rates can happen when the rates fluctuate too much, especially when the number of substitutions is low (Sanderson, 2002), as is the case in our data. Thus, the best model according to the cross-validation analysis was equivalent to a molecular clock.

Inferred ages for the South American clades—The Langley–Fitch algorithm was used in r8s (Sanderson, 2003) to infer ages for the two South American crown nodes using the estimated age of the Neo-Astragalus clade (Wojciechowski et al., 2005) as a calibration point. Clade F had a mean crown age of 1.89 ± 0.18 Ma and clade G, a mean crown age of 0.98 ± 0.19 Ma (Table 1). An independent analysis of the ages of the clades enforcing a molecular clock was done using PAUP*. Almost identical ages for clades F and G were obtained (Table 1).

Rates of species diversification for the South American clades—The two South American clades had almost the same rate of diversification of 2.01 ± 0.07 spp./Ma for clade F and 2.07 ± 0.14 spp./Ma for clade G, when the crown node age was used. The upper and lower bounds for the rates when crown vs. stem ages and sampled vs. hypothesized number of species were used, changed considerably more for clade F than for clade G (Table 1). This change of rates in clade F was a consequence of the differences between the crown and stem ages and between

the hypothesized vs. sampled number of species, which are more pronounced in this clade than in clade G.

Phylogenetic analyses—Separate ITS or chloroplast data sets yielded trees with only moderate support. The ITS tree recovered the Neo-Astragalus clade with high support and *A. adsurgens* as its sister taxon. However, within Neo-Astragalus, the resolution was poor. The chloroplast spacers also recovered the Neo-Astragalus clade with high support, but its sister taxon was not clear due to lack of resolution at the base of the tree. Both trees had two disjoint South American clades within Neo-Astragalus (Appendix S4, see Supplemental Data with the online version of this article).

A dramatic increase in clade support was obtained when the ITS and chloroplast data sets were combined, either in a partitioned or a nonpartitioned Bayesian approach. The trees were highly similar, with the exception of the placement of *A. calycosus* Torr. ex S. Watson and *A. mollissimus* Torr., which were either nested within the largest clade (A in Fig. 2) in the nonpartitioned tree or formed a polytomy at the base of this clade in the partitioned approach. Both trees had two main clades within Neo-Astragalus, (A and B in Fig. 2). Clade A had the most species and included one South American clade (clade G) nested within it. Clade B was formed by a highly supported group of species from South America (clade F) as a sister group of the clade formed by the North American species *A. nothoxys* A. Gray and *A. arizonicus* A. Gray.

A maximum likelihood bootstrap analysis using the combined data set had improved resolution when compared to the separate analyses and also supported the major clades obtained in the combined Bayesian trees. However, the resolution of some of the clades, especially those close to the root, was much lower than with the Bayesian approach (Fig. 2). This difference in clade support may be a reflection of the generally higher posterior probabilities relative to bootstrap values (Alfaro and Holder, 2006).

DISCUSSION

High rates of species diversification: A general trend?—Even the lowest estimates of rates of diversification in the South American clades of Neo-Astragalus are very high. Until recently, the most spectacular cited examples of rapid speciation were radiations in island or island-like environments. Classic examples such as the Darwin finches (Sato et al., 1999; Grant and Grant, 2005) and the Hawaiian Silversword alliance with reported rates of 0.56 spp/Ma (Baldwin and Sanderson, 1998) have been linked to adaptation to newly available ecological niches, often taking advantage of key morphological innovations (Schluter, 2000). However, recent studies are showing comparable or higher rates of species diversification in nonisland taxa (Table 2).

For Neo-Astragalus as a whole, using the estimated age of the crown node of 4.4 Ma and an approximate number of 500 species (Barneby, 1964; ILDIS, 2002; Gómez-Sosa, 2005), the rate of diversification can be calculated as 1.25 spp./Ma, very close to previous estimates of 1.48 spp./Ma (Wojciechowski et al., 1999). The calculated rates of diversification of both South American clades are even higher, 2.01 and 2.07 spp./Ma for clades F and G, respectively. Comparably high rates of species diversification (Table 2) have been reported in many taxa in the Andes, which is an area of high endemism. Examples include *Valeriana* L., with multiple invasions from North to South America in the last 5 Ma (Bell and Donoghue, 2005) or the Andean clade of *Lupinus,* endemic to the tropical high Andes (Hughes and Eastwood, 2006). The Iochrominae (Solanaceae) (Smith and Baum, 2006) and *Schizanthus* Ruiz & Pav. (Perez et al., 2006), both endemic to the Andes, experienced rapid speciation associated with diversification of flower morphology and pollination biology. In the neotropical genus *Costus* L. (Costaceae), high rates have also been associated with hummingbird pollination syndromes (Kay et al., 2005). The tribe Vaccineae (Ericaceae), inhabiting middle elevations is especially diverse, with high levels of endemism (Luteyn, 2002). The Espeletiinae (Asteraceae, Heliantheae), endemic to the tropical high Andes with very recent origin has remarkable morphological diversity adapted to many environments (Rauscher, 2002). Ecological opportunities provided by the different environments and microclimates available after the uplift of the Andes in the Pliocene, 2–4 Ma (Van der Hammen, 1989) are believed to influence this observed diversity (Ortiz-Jaureguizar and Cladera, 2006).

Use of partitions in a Bayesian framework—Some recent studies have used different evolutionary models for analyzing partitions of a single data set in a Bayesian framework (Brandley et al., 2005; Strugnell et al., 2005; Clayton et al., 2007). In cephalopods, Strugnell et al. (2005) found better resolution and support for clades when partitions were used compared to a single model. The same result was observed by Brandley et al. (2005) when comparing partition choices in scincid lizards using the mitochondrial ND1 gene. However, for the DNA regions used in this study, the partitioned approach differed very little in topology and clade support compared to the nonpartitioned analysis. It is possible that partitions have more effect when coding genes are used. In the studies mentioned, more support was obtained for trees in which the data were partitioned by gene +

TABLE 1.   Inferred mean ages and rates of species diversification with standard deviations of the posterior distribution of the two South American clades of Neo-Astragalus using upper and lower bounds of ages and numbers of species. Standard deviations for the posterior distributions are given for all values. Standard deviation for ages was calculated considering the error from the source of calibration (see Materials and Methods). The standard deviations in rates are only due to the analyses in this study.

Clade	Inferred no. species	Mean crown group age (Ma)	Mean stem group age (Ma)	Mean crown group rate with N (spp/Ma)	Mean stem group rate with N (spp/Ma)	Mean crown group rate with S (spp/Ma)	Mean stem group rate with S (spp/Ma)
F	90	(a) 1.89 ± 0.18* (b) 1.80 ± 0.18*	3.38 ± 0.18	2.01 ± 0.07	1.33 ± 0.04	0.79 ± 0.03	0.65 ± 0.02
G	15	(a) 0.98 ± 0.19* (b) 1.04 ± 0.19*	1.68 ± 0.19	2.07 ± 0.14	1.62 ± 0.11	1.42 ± 0.10	1.24 ± 0.08

Notes: (a) Result obtained using r8s, (b) result obtained using PAUP*. * Values for clade F and clade G are significantly different with $P < 0.0001$ (*t* test). Ma = million years, N = number of species hypothesized to exist in the clade, S = number of species sampled, spp/Ma = species per million years.
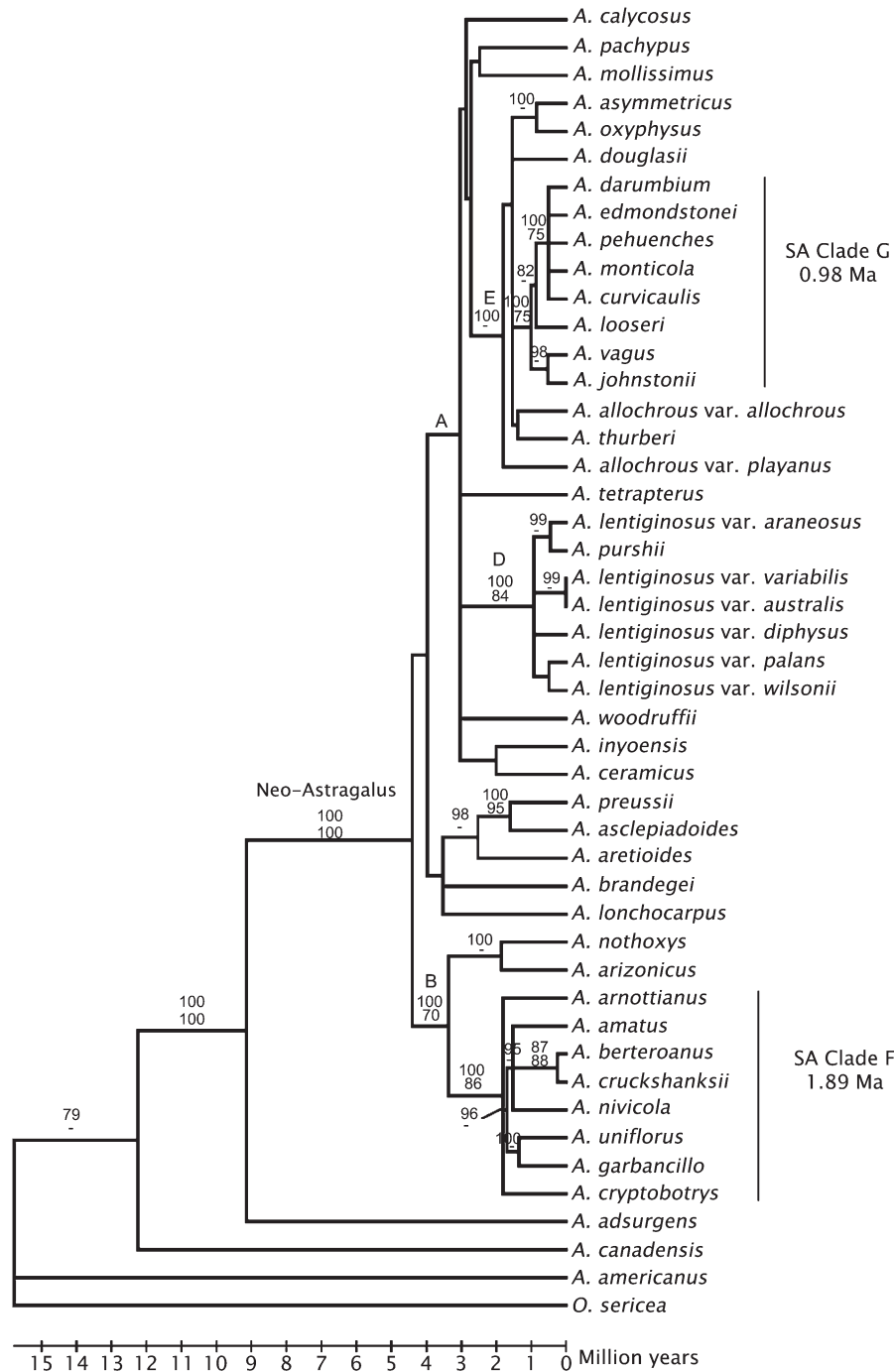
Fig. 2. Chronogram depicting relationships among species of Neo-Astragalus. The topology of the tree was obtained using Bayesian analyses with both data sets (ITS and the chloroplast spacers *trnD-trnT* and *trnfM-trnS1*) as one matrix, implementing the GTR+I+G model without partitions. The numbers represent statistical support for the clades. Numbers above are posterior probabilities, and the numbers below are bootstrap percentages, as obtained in the maximum likelihood analysis. The largest well-supported clades are also labeled with letters. South American (SA) clades with their corresponding calculated ages are indicated with a vertical line. The rest correspond to North American species, except for *A. adsurgens* which is also present in Asia. Outgroup is *Oxytropis sericea*.

codon or codon alone. Such partitions are obviously not available for noncoding sequences. Clayton et al. (2007) used both coding and noncoding sequences in the phylogeny of Simaroubaceae (Sapindales). Even though they did observe a significant difference in the likelihood scores when the data were fully partitioned, this difference did not affect the topology of the trees.

Despite this result, when using the same complex model in a maximum likelihood analysis with the combined data set, the support for some clades was lower. The difference in support for clades when using Bayesian analyses as compared to nonparametric bootstrapping has been extensively discussed (Wilcox et al., 2002; Alfaro et al., 2003; Erixon et al., 2003; Lemmon and Moriarty,

TABLE 2. Some recent studies of taxa with high reported rates of species diversification.

Taxon	Geographic area	Rate (spp./Ma)	Reference
Aizoaceae	South Africa	0.77–1.45	Klak et al., 2004
Gentaniella Moench	Australia, NZ	1.73	von Hagen and Kadereit, 2001
	South America	1.48–1.71	
Valeriana L.	Andean Paramo	1.71–3.2	Bell and Donoghue, 2005
Lupinus L.	Tropical high Andes	2.49–3.72	Hughes and Eastwood, 2006
Costus L.	Neotropical Central and Northern South America	0.6–2.6	Kay et al., 2005
Neo-Astragalus	North and South America	1.48	Wojciechowski et al., 1999
Woodland salamanders (*Plethodon* sp.)	North America	0.8	Kozak et al., 2006
Logperch darters (*Percina* sp.)	Eastern North America	0.96	Near and Benard, 2004

2004; Taylor and Piel, 2004; Alfaro and Holder, 2006) and is also evidenced in this study. Even though more analyses would be necessary to assess the accuracy of both approaches, we did not consider the need for additional analyses because no clade conflict was observed among trees. In addition, all the important clades are well supported in both the Bayesian and the ML analyses, despite the lack of resolution at the base of some clades in the maximum likelihood tree.

Phylogeny of Neo-Astragalus—The utility of the ITS in phylogenetic analyses has occasionally been questioned because of the possibility that paralogous regions and/or pseudogenes exist within repeats of the rDNA arrays (Alvarez and Wendel, 2003; Li et al., 2004). However, ITS has been instrumental in reconstructing phylogenies at low taxonomic levels, especially in combination with chloroplast spacers or other nuclear genes (Lavin et al., 2001; McMahon and Hufford, 2004; Kenicer et al., 2005). As previously shown (Scherson et al., 2005), the presence of paralogous sequences can be evidenced by more than one band on an agarose or polyacrylamide gel and in paralogous sequences of the same size, by studying the sequence chromatographs. In this study, there was no evidence of paralogous divergence in any of the taxa used.

Neo-Astragalus was recovered with high support in all the analyses, which reaffirms previous studies (Wojciechowski et al., 1993, 1999; Sanderson and Wojciechowski, 2000). Two distinct South American clades nested within Neo-Astragalus are also very well supported. However, this support needs to be taken with caution since incomplete sampling might render higher values to clades where sampling is sparse (Lecointre et al., 1993; Sanderson and Wojciechowski, 2000).

Wojciechowski et al. (1999) remarked on the correspondence of the two South American clades with the two sections proposed by Johnston (1947) based on gross morphology. The present work provides further evidence for this correspondence. Except for *A. berteroanus* Moris (Reiche), all species of clade G have herbaceous and free stipules, whereas the species in clade F have membranous and more or less united stipules. Interestingly, and despite their different stipule morphology, Johnston (1947) suggested a very close relationship between *A. berteroanus* and *A. looseri* I. M. Johnst., which have very similar habit, shape, and structure of fruit. The present phylogenies do not support this relationship, being *A. berteroanus* within clade F and *A. looseri* in clade G. It is odd then that *A. berteroanus* is the only representative of clade F with herbaceous free stipules and is morphologically similar to species in clade G. The initial division by Johnston in 1947 did not, of course, include the recently discovered species *A. johnstonii* Gómez-Sosa

(Gómez-Sosa, 1997), endemic to a very narrow area in the coastal ranges of central Chile. This species is described as having membranous and free stipules, which is unique within the South American species of the genus described by Johnston (1947). The fact that *A. johnstonii* is well nested within clade G, in which the rest of the species have herbaceous and free stipules, suggests that the degree of union of the stipules might be more diagnostic character than their texture.

A second migration from North America would have given rise to the South American clade F, whose sister taxa are the North American species *A. nothoxys* and *A. arizonicus*, both representatives of section *Leptocarpi* (Barneby, 1964). The latter species have an overlapping geographical distribution in Arizona and northern Mexico. Other nonsampled species in this section occupy a larger geographic area, in southern California, Texas, and southern Mexico (Barneby, 1964). Because closest outgroups to the South American species could be found within these southern Mexican species, further sampling of this section is obviously desirable.

Biogeography of Neo-Astragalus—Even though the highest center of diversity in *Astragalus* is in central Asia, this and previous studies (Wojciechowski et al., 1999; Wojciechowski, 2005) suggest that the South American clades of Neo-Astragalus came from North America. This pattern was earlier noted by Johnston (1947), stating that the species in South America are much more similar morphologically to the ones found in North America than to the ones found in central Asia. The vicariant model, in which there was one widespread *Astragalus* that split into North and South American species more recently than did the New World from the Old World, cannot be ruled out. However, we believe this model to be unlikely because the South American species are nested with strong statistical support within the North American ones, and based on molecular clock analyses, the clades appear younger than the North American groups. In addition, the vicariant model in this case would have entailed many episodes of extinction. The narrow geographical ranges covered by the majority of the species of *Astragalus*, and the recent ages calculated for the clades also argue against this hypothesis.

The South American species in clade G (Fig. 2) appear closely related to species of Barneby's section *Inflati*, distributed in southwestern United States and adjoining Mexico and to what Barneby called the Pacific Piptolobi, in the coast of southern and Baja California (Barneby, 1964). Despite being separated by more than 7500 km, the South and North American species are morphologically similar, with almost indistinguishably different large (~3–5 cm), unilocular, inflated pods. Wojciechowski et al. (1999) noted this pattern and suggested the possibility of multiple

migrations to South America or back to North America. However, in our analyses the South American species form a well-supported clade nested within the wider North American clade G. This pattern rather suggests a single colonization of South America possibly originating from California or Baja California with subsequent in situ diversification in the Andes. This result seems to refute the idea of a general pattern of amphitropical disjunction at the level of sister species, even though a more comprehensive sampling is needed to estimate patterns of disjunction of sister clades.

There are important differences between clade F and clade G regarding their geographical distribution (Fig. 3). Species in clade G are concentrated in central Chile and Argentina and generally do not reach high elevations, whereas species in clade F which have a wider latitudinal distribution, are found from Ecuador to Patagonia, and reach higher altitudes. It is interesting that in contrast to what is seen in other Andean plant taxa in which the highest diversity is in central and northern Andes, species diversity in *Astragalus* decreases as it moves north. In fact, there are no species of *Astragalus* reported north of Ecuador, even though a few specimens from *A. garbancillo* have

been found in Colombia (Tropicos.org, 2008). This distribution mirrors in some ways the pattern seen in North America, where *Astragalus* richness drops dramatically between arid temperate regions (e.g., the Colorado Plateau in the U.S.) and more subtropical areas in Mexico and further south, regions where diversity in many other legume taxa skyrockets.

Although our new sample of phylogenetic history in *Astragalus* seems to reveal several important patterns, many detailed questions regarding the biogeographic history of the South American radiations of Neo-Astragalus may have to await a nearly complete phylogeny that encompasses all of its complex history.

LITERATURE CITED

ALFARO, M. E., AND M. T. HOLDER. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology Evolution and Systematics* 37: 19–42.

ALFARO, M. E., S. ZOLLER, AND F. LUTZONI. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 20: 255–256.

ALLAN, G. J., J. FRANCISCO-ORTEGA, A. SANTOS-GUERRA, E. BOERNER, AND E. A. ZIMMER. 2004. Molecular phylogenetic evidence for the geographic origin and classification of Canary Island lotus (Fabaceae: Loteae). *Molecular Phylogenetics and Evolution* 32: 123–138.

ÁLVAREZ, I., AND J. F. WENDEL. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417–434.

BALDEÓN, S., M. FLORES, AND J. ROQUE. 2006. Fabaceae endémicas del Perú. *Revista Peruana de Biologia* 13: 302s–337s.

BALDWIN, B. G., AND M. J. SANDERSON. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences, USA* 95: 9402–9406.

BARNEBY, R. C. 1964. Atlas of North American *Astragalus. Memoirs of the New York Botanical Gardens* 13: 1–1188.

BELL, C. D., AND M. J. DONOGHUE. 2005. Phylogeny and biogeography of Valerianaceae (Dipsacales) with special reference to the South American valerians. *Organisms, Diversity & Evolution* 5: 147–159.

BRANDLEY, M. C., A. SCHMITZ, AND T. W. REEDER. 2005. Partitioned Bayesian analyses, partition choice and the phylogenetic relationships of scincid lizards. *Systematic Biology* 54: 373–390.

CLAYTON, J. W., E. S. FERNANDO, P. S. SOLTIS, AND D. E. SOLTIS. 2007. Molecular phylogeny of the tree-of-heaven family (Simaroubaceae) based on chloroplast and nuclear markers. *International Journal of Plant Sciences* 168: 1325–1339.

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.

GALTIER, N., M. GOUY, AND C. GAUTIER. 1996. SEAVIEW and PHYLO_WIN: Two graphic tools for sequence alignment and molecular phylogeny. *Computer Applications in the Biosciences* 12: 543–548.

GÓMEZ-SOSA, E. 1979. Las especies sudamericanas del género *Astragalus* (*Leguminosae*). I. Las especies patagónicas argentinas. *Darwiniana* 22: 313–376.

GÓMEZ-SOSA, E. 1994. *Astragalus* Linnaeus. *In* R. Kiesling [ed.], Flora de San Juan. Vazquez Maxxini Editores, Buenos Aires, Argentina.

GÓMEZ-SOSA, E. 1997. *Astragalus johnstonii* sp. nov. (Fabaceae) y relaciones con el complejo *A. verticillatus* (Phil.) Reiche. *Gayana Botánica* 54: 31–37.

GÓMEZ-SOSA, E. 2003. El género *Astragalus* (Leguminosae) en la república oriental del Uruguay. *Hickenia 3* 40: 161–162.

GÓMEZ-SOSA, E. 2005. Taxonomic novelties in *Astragalus* (Leguminosae) for South America. *Novon* 15: 542–547.

GOODMAN, L. A. 1960. On the exact variance of products. *Journal of the American Statistical Association* 55: 708–713.

GRANT, P. R., AND B. R. GRANT. 2005. Darwin's finches. *Current Biology* 15: R614–R615.

Fig. 3. Geographic distribution pattern of the South American taxa of Neo-Astragalus sampled in this study. Stars represent the general distribution of the sampled species in clade G, and circles represent the pattern of distribution of sampled species in clade F. Whereas species in clade G are concentrated in central Chile and Argentina (and generally at lower elevations—not shown), species in clade F are more widely distributed, both in latitude as well as in elevation.

HUGHES, C., AND R. EASTWOOD. 2006. Island radiation on a continental scale: Exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences, USA* 103: 10334–10339.

ILDIS [International Legume Database & Information Service]. 2002. World database of legumes. ILDIS, Southampton, UK.

JOHNSTON, I. M. 1947. *Astragalus* in Argentina, Bolivia and Chile. *Journal of the Arnold Arboretum. Arnold Arboretum* 28: 336–409.

KAY, K. M., P. A. REEVES, R. G. OLMSTEAD, AND D. W. SCHEMSKE. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. *American Journal of Botany* 92: 1899–1910.

KENICER, G. J., T. KAJITA, R. T. PENNINGTON, AND J. MURATA. 2005. Systematics and biogeography of *Lathyrus* (Leguminosae) based on internal transcribed spacer and cpDNA sequence data. *American Journal of Botany* 92: 1199–1209.

KLAK, C., G. REEVES, AND T. HEDDERSON. 2004. Unmatched tempo of evolution in southern African semi-desert ice plants. *Nature* 427: 63–65.

LAVIN, M., P. S. HERENDEEN, AND M. F. WOJCIECHOWSKI. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology* 54: 575–594.

LAVIN, M., R. T. PENNINGTON, B. B. KLITGAARD, J. I. SPRENT, H. C. DE LIMA, AND P. E. GASSON. 2001. The dalbergioid legumes (Fabaceae): Delimitation of a pantropical monophyletic clade. *American Journal of Botany* 88: 503–533.

LECOINTRE, G., H. PHILIPPE, H. L. V. LE, AND H. LE GUYADER. 1993. Species sampling has a major impact on phylogenetic inference. *Molecular Phylogenetics and Evolution* 2: 205–224.

LEMMON, A. R., AND E. C. MORIARTY. 2004. The importance of proper model assumption in Bayesian phylogenetics. *Systematic Biology* 53: 265–277.

LI, D. Y., Y. Y. RU, AND X. Y. ZHANG. 2004. Chromosomal distribution of the 18S-5.8S-26S rDNA loci and heterogeneity of nuclear in regions in *Thinopyrum intermedium* (Poaceae: Triticeae). *Acta Botanica Sinica* 46: 1234–1241.

LINDER, H. P., C. R. HARDY, AND F. RUTSCHMANN. 2005. Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution* 35: 569–582.

LISTON, A. 1989. Why are there so many species of *Astragalus? American Journal of Botany* 76 (supplement): 256 (abstract).

LOCK, J. M., AND K. SIMPSON. 1991. Legumes of West Asia, a check list. Royal Botanical Gardens, Kew, UK.

LUTEYN, J. L. 2002. Diversity, adaptation, and endemism in neotropical Ericaceae: Biogeographical patterns in the Vaccinieae. *Botanical Review* 68: 55–87.

MABBERLEY, D. J. 1997. The plant book. Cambridge Univeristy Press, Cambridge, UK.

MAGALLÓN, S., AND M. J. SANDERSON. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55: 1762–1780.

MASSOUMI, A. A. R. 1998. Old World check-list of *Astragalus*. Research Institute of Forests and Rangelands, Tehran, Iran.

McMAHON, M., AND L. HUFFORD. 2004. Phylogeny of Amorpheae (Fabaceae: Papilionoideae). *American Journal of Botany* 91: 1219–1230.

NEAR, T. J., AND M. F. BENARD. 2004. Rapid allopatric speciation in logperch darters (Percidae: *Percina*). *Evolution* 58: 2798–2808.

NYLANDER, J. A. A. 2002. MrModeltest 2. Computer program distributed by the author, website http://www.ebc.uu.se/systzoo/staff/nylender.html.

ORTIZ-JAUREGUIZAR, E., AND G. A. CLADERA. 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. *Journal of Arid Environments* 66: 498–532.

OSALOO, S. K., A. A. MAASSOUMI, AND N. MURAKAMI. 2005. Molecular systematics of the Old World Astragalus (Fabaceae) as inferred from nrDNA ITS sequence data. *Brittonia* 57: 367–381.

PÉREZ, F., M. T. K. ARROYO, R. MEDEL, AND M. A. HERSHKOVITZ. 2006. Ancestral reconstruction of flower morphology and pollination systems in *Schizanthus* (Solanaceae). *American Journal of Botany* 93: 1029–1038.

POSADA, D., AND T. BUCKLEY. 2004. Model selection and model averaging in phylognetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53: 793–808.

RAMBAUT, A. 1996. Se-Al: Sequence Alignment Editor. University of Oxford, Oxford, UK.

RAUSCHER, J. T. 2002. Molecular phylogenetics of the *Espeletia* complex (Asteraceae): Evidence from nrDNA its sequences on the closest relatives of an Andean adaptive radiation. *American Journal of Botany* 89: 1074–1084.

RICHARDSON, J. E., R. T. PENNINGTON, T. D. PENNINGTON, AND P. M. HOLLINGSWORTH. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* 293: 2242–2245.

RONQUIST, F., AND J. P. HUELSENBECK. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* 19: 1572–1574.

SANDERSON, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19: 101–109.

SANDERSON, M. J. 2003. r8s: Inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics (Oxford, England)* 19: 301–302.

SANDERSON, M. J., AND M. F. WOJCIECHOWSKI. 1996. Diversification rates in a temperate legume clade: Are there "so many species" of *Astragalus* (Fabaceae)? *American Journal of Botany* 83: 1488–1502.

SANDERSON, M. J., AND M. F. WOJCIECHOWSKI. 2000. Improved bootstrap confidence limits in large-scale phylogenies, with an example from Neo-Astragalus (leguminosae). *Systematic Biology* 49: 671–685.

SATO, A., C. O'HUIGIN, F. FIGUEROA, P. R. GRANT, B. R. GRANT, H. TICHY, AND J. KLEIN. 1999. Phylogeny of Darwin's finches as revealed by mtDNA sequences. *Proceedings of the National Academy of Sciences, USA* 96: 5101–5106.

SCHERSON, R. A., H. K. CHOI, D. R. COOK, AND M. J. SANDERSON. 2005. Phylogenetics of New World *Astragalus*: Screening of novel nuclear loci for the reconstruction of phylogenies at low taxonomic levels. *Brittonia* 57: 354–366.

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

SHAW, J., E. B. LICKEY, J. T. BECK, S. B. FARMER, W. S. LIU, J. MILLER, K. C. SIRIPUN, C. T. WINDER, E. E. SCHILLING, AND R. L. SMALL. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* 92: 142–166.

SMITH, S. D., AND D. A. BAUM. 2006. Phylogenetics of the florally diverse Andean clade Iochrominae (Solanaceae). *American Journal of Botany* 93: 1140–1153.

STRUGNELL, J., M. NORMAN, J. JACKSON, A. J. DRUMMOND, AND A. COOPER. 2005. Molecular phylogeny of coleoid cephalopods (Mollusca: Cephalopoda) using a multigene approach; the effect of data partitioning on resolving phylogenies in a Bayesian framework. *Molecular Phylogenetics and Evolution* 37: 426–441.

SWOFFORD, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Massachusetts, USA.

TAYLOR, D. J., AND W. H. PIEL. 2004. An assessment of accuracy, error, and conflict with support values from genome-scale phylogenetic data. *Molecular Biology and Evolution* 21: 1534–1537.

Tropicos.org. 2008. Missouri Botanical Garden, St. Louis, Missouri, USA. Website http://www.tropicos.org.

VAN DER HAMMEN, T. 1989. History of the mountain forest of the northern Andes. *Plant Systematics and Evolution* 162: 109–114.

WHITE, T. J., T. D. BRUNS, S. B. LEE, AND J. W. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* M. A. Innes, D. H. Gelfand, J. J. Sninsky, and T. J. White [eds.], PCR protocols: A guide to methods and applications, 315–322. Academic Press, San Diego, California, USA.

WILCOX, T. P., D. J. ZWICKL, T. A. HEATH, AND D. M. HILLIS. 2002. Phylogenetic relationships of the dwarf boas and a compari-

son of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution* 25: 361–371.

WOJCIECHOWSKI, M. F. 2005. Astragalus (Fabaceae): A molecular phylogenetic perspective. *Brittonia* 57: 382–396.

WOJCIECHOWSKI, M. F., M. J. SANDERSON, B. G. BALDWIN, AND M. J. DONOGHUE. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): Evidence from nuclear ribosomal DNA internal transcribed spacer sequences. *American Journal of Botany* 80: 711–722.

WOJCIECHOWSKI, M. F., M. J. SANDERSON, AND J. M. HU. 1999. Evidence on the monophyly of *Astragalus* (Fabaceae) and its major subgroups based on nuclear ribosomal DNA ITS and chloroplast DNA *trnL* intron data. *Systematic Botany* 24: 409–437.

WOJCIECHOWSKI, M. F., K. P. STEEL, AND M. J. SANDERSON. 2005. Are there so many species of Astragalus? A reprise. XVII International Botanical Congress, Vienna Austria, 11 [online abstract http://www.ibc2005.ac.at/].