Molecular Phylogenetics and Evolution 65 (2012) 621–630

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Species tree estimation for a deep phylogenetic divergence in the New World monkeys (Primates: Platyrrhini)

S. Ivan Perez [a,*], Julia Klaczko [b], Sérgio F. dos Reis [b]

[a] División Antropología, Museo de La Plata, Universidad Nacional de La Plata, CONICET, Paseo del Bosque s/n, 1900 La Plata, Argentina
[b] Departamento de Biologia Animal, Universidade Estadual de Campinas, 13083-970 Campinas, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 3 December 2011
Revised 13 June 2012
Accepted 16 July 2012
Available online 24 July 2012

Keywords:
Gene tree
Species tree
*BEAST
Molecular datasets
Platyrrhines
America

ABSTRACT

The estimation of a robust phylogeny is a necessary first step in understanding the biological diversification of the platyrrhines. Although the most recent phylogenies are generally robust, they differ from one another in the relationship between *Aotus* and other genera as well as in the relationship between Pitheciidae and other families. Here, we used coding and non-coding sequences to infer the species tree and embedded gene trees of the platyrrhine genera using the Bayesian Markov chain Monte Carlo method for the multispecies coalescent (*BEAST) for the first time and to compared the results with those of a Bayesian concatenated phylogenetic analysis. Our species tree, based on all available sequences, shows a closer phylogenetic relationship between Atelidae and Cebidae and a closer relationship between *Aotus* and the Cebidae clade. The posterior probabilities are lower for these conflictive tree nodes compared to those in the concatenated analysis; this finding could be explained by some gene trees showing no concordant topologies between *Aotus* and the other genera. Moreover, the topology of our species tree also differs from the findings of previous molecular and morphological studies regarding the position of *Aotus*. The existence of discrepancies between morphological data, gene trees and the species tree is widely reported and can be related to processes such as incomplete lineage sorting or selection. Although these processes are common in species trees with low divergence, they can also occur in species trees with deep and rapid divergence. The sources of the inconsistency of morphological and molecular traits with the species tree could be a main focus of further research on platyrrhines.

1. Introduction

The New World monkeys, constituting the parvorder Platyrrhini, are a monophyletic clade with a history of nearly 25–35 million years in South America and the Caribbean (Fleagle, 1999; Tejedor, 2008; Wilkinson et al., 2011). They invaded the continent during the Paleocene–Oligocene and evolved in isolation from the Old World primates (Rosenberger, 2002; Tejedor, 2008; Opazo et al., 2006). In America, they experienced a great biological radiation, resulting in the diversification of the lineages that represent the majority of the extant 15–17 genera and 125 species, occupation of a large range of ecological niches and great morphological variation (Rosenberger, 1992; Fleagle, 1999; Norconk et al., 2009; Wildman et al., 2009; Perez et al., 2011).

The biological diversification of the platyrrhine clade and of several orders and suborders of Mammalia is peculiar because it began in the distant past and then continued as a hierarchically structured process during phylogenetic divergence (Rosenberger, 2002). Because of the long history of platyrrhine radiation, the estimation of a robust phylogeny that is relatively independent of other biological data is a necessary first step for understanding the factors responsible for the biological diversification of this clade (Wiens, 2009; Freckleton et al., 2011; Losos, 2011). Researchers who have studied the branching pattern or species tree of the platyrrhines using morphological data have obtained different results (Rosenberger, 1984; Ford, 1986; Kay, 1990). The most recent phylogenetic studies have mostly analyzed DNA sequences and/or a concatenation of these sequences and have presented different views of the pattern of divergence among platyrrhine genera (e.g., Goodman et al., 1998; Horovitz et al., 1998; Schneider, 2000; Opazo et al., 2006; Wildman et al., 2009; Perelman et al., 2011). Although these most recent phylogenies are generally robust, they differ in some of the topological relationships among genera and families. In particular, these studies have presented topologies that mainly differ from each other in the relationship between *Aotus* and other genera (e.g., Wildman et al., 2009; Perelman et al., 2011) and the relationships between Pitheciidae and other families (e.g., Schneider et al., 2001; Opazo et al., 2006).

There could be several explanations for these differences; the most probable explanation is that the gene analyzed can have an

* Corresponding author. Fax: +54 2214257744.
E-mail addresses: iperez@fcnym.unlp.edu.ar, ivanperezmorea@gmail.com (S.I. Perez).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.07.014

evolutionary history that is distinct from the underlying species tree due to processes such as incomplete lineage sorting (Maddison, 1997; Rannala and Yang, 2008; Degnan and Rosenberg, 2009; Heled and Drummond, 2010; Knowles and Kubatko, 2010). The approaches employed previously (i.e., gene tree and concatenated analyses; e.g., Horovitz et al., 1998; Schneider, 2000; Wildman et al., 2009) may not allow a species tree to be estimated if the individual genes examined have different topologies (i.e., branching patterns; Liu et al., 2008; Degnan and Rosenberg, 2009; Knowles and Kubatko, 2010), which could be true for the platyrrhine clade (see below). Moreover, previous works have found that when there are high levels of discordance among gene trees the concatenation method can result in an incorrect species tree as more data are added (Degnan and Rosenberg, 2009). A Bayesian method, which is based on the multispecies coalescent model, has recently been proposed to co-estimate multiple gene trees embedded in a shared species tree from multiple-gene data (Liu, 2008; Heled and Drummond, 2010). This method allows and uses high levels of gene tree discordance when inferring species trees (Liu et al., 2008; Degnan and Rosenberg, 2009; Heled and Drummond, 2010; Leaché and Rannala 2011). Obtaining a species tree based on coalescent theory is a necessary next step for resolving platyrrhine systematics.

Here, we study the phylogenetic or species tree (i.e., the pattern and relative timing of divergence) of the New World monkeys using molecular data and discuss the differences compared to previously published phylogenetic trees. Specifically, we first estimate the phylogenetic tree of the platyrrhine genera based on a multispecies coalescent model and all of the available datasets. Second, we explore whether the differences observed in the topological relationships among platyrrhine genera and families between our species tree and the previously reported phylogenetic trees could be related to the use of particular gene sequences and the most popular concatenated methods. To explore these questions, we use several datasets containing multiple coding and non-coding DNA sequences (Opazo et al., 2006; Wildman et al., 2009; Perelman et al., 2011), infer the species tree and embedded gene trees of the platyrrhine genera from multiple gene sequences using the Bayesian Markov chain Monte Carlo method for the multispecies coalescent for the first time (Liu, 2008; Heled and Drummond, 2010) and compare these results with the phylogenetic tree based on a partitioned Bayesian analysis of concatenated sequences (Ronquist and Huelsenbeck, 2003; Drummond and Rambaut, 2007).

2. Materials and methods

2.1. Molecular datasets

To study the evolutionary history of the New World monkey genera, we analyzed 15–17 members of the platyrrhines and several outgroups (*Macaca mulatta*, *Pan troglodytes*, *Homo sapiens*). Three different molecular datasets were obtained from GenBank and from the supporting information in the original studies (Wildman et al., 2009; Perelman et al., 2011). Specifically, we used the post-GBLOCK editing alignment made available by Perelman et al. (2011) and the alignment data made available by Wildman et al. (2009). Sequences obtained from GenBank were aligned using ClustalW, and the alignment was manually corrected and the ambiguous regions removed with BioEdit 7.0.0 software (Hall, 2004).

First, we combined the datasets to perform a main phylogenetic analysis using all of the available data including 68 loci (47,233 bp) from 15 genera. We also re-analyzed all of the available data for 15 genera excluding missing loci (27,360 bp). We refer to this dataset as the combined dataset. To explore the differences between our main estimation of the platyrrhine species tree based on all available data and previous estimations, we performed phylogenetic analyses using different sub-datasets corresponding to the datasets used in previous studies. The first sub-dataset analyzed is a molecular matrix of four nuclear coding gene sequences from 15 genera, constituting a 5392 bp matrix (Table 1). These sequences were previously used by Opazo et al. (2006). The second sub-dataset comprises 10 non-coding sequences, obtained by Wildman and co-workers (2009) from many different chromosomes of 15 genera, and constitutes a 6899 bp matrix (Table 1). The third sub-dataset is based on 54 coding and non-coding sequences obtained by Perelman and co-workers (2011); it comprises 68 species, 17 genera and a 34,941 bp matrix. Please note that for the Perelman sub-dataset, we also generated a summarized dataset choosing only the genes that were sequenced for all of the selected genera in an attempt to minimize the missing data. This dataset is based on 22 coding and non-coding sequences; it comprises 17 genera and constitutes a 15,068 bp matrix (Table 1). See accession numbers in the cited works (Opazo et al., 2006; Wildman et al., 2009; Perelman et al., 2011).

2.2. Gene tree and species tree estimations

jModelTest 0.1.1 (Posada, 2008) was employed to determine the most appropriate model of sequence evolution for each analyzed gene estimated under the Akaike Information Criterion with correction for sample size (AICc). The best-fitting model for each sequence studied is shown in Table 1 and Supplementary Table A.1. Models of sequence evolution identified as optimal by jModelTest for both coding and non-coding sequences were implemented in the following analyses.

The species and gene trees for each dataset were estimated jointly using ∗BEAST (Heled and Drummond, 2010), which is part of the BEAST v1.6.1 package (Drummond and Rambaut, 2007). ∗BEAST uses the multispecies coalescent, an extension of the prior coalescent designed to handle multiple species. Each gene tree is embedded in a shared species tree and follows the coalescent in each extant and ancestral species. Please note that the term "species" used in this work is not the same as the taxonomic rank and instead designates a group of individuals that likely have no history of breeding with individuals outside of that group (Heled and Drummond, 2010). We use the BEAUti program to unlink the substitution models of the data partitions and to implement the models of sequence evolution identified as optimal by jModelTest. We set the clock model to the strict clock model and the species tree priors as a Yule Process as suggested by Heled and co-workers (2011). Three simultaneous analyses were performed using Markov Chain Monte Carlo (MCMC) simulations for 400,000,000–200,000,000 generations, with a sampling frequency of 40,000–20,000 in the program BEAST. The convergence of the analyses was determined with the program Tracer v1.5 (Rambaut and Drummond, 2007). Finally, we discarded the first 1250 trees as burn-in and summarized the trees using TreeAnnotator v1.4.8 (Drummond and Rambaut, 2007).

For comparative purposes, we also estimated the phylogenetic tree using concatenated Bayesian analyses based on the datasets. Models of sequence evolution identified as optimal for the data partitions by jModelTest were implemented for each of the data partitions. These phylogenetic analyses were performed using BEAST v1.6.1 (Drummond and Rambaut, 2007). Three simultaneous analyses were performed using MCMC simulations for 200,000,000 generations with a sampling frequency of 20,000. The convergence was determined with Tracer v1.5 (Rambaut and Drummond, 2007) and the first 1250 trees sampled were excluded. We analyzed the sequences under a relaxed molecular clock model, which allows

Table 1
Coding and non-coding sequences used in the current study, sequence size (bp), substitution models and gene tree results. See accession numbers in Opazo et al. (2006), Wildman et al. (2009) and Perelman et al. (2011).

Dataset	Sequence	Sequence type	bp	Substitution model	nst	Rates	Gene tree in the Supplementary Fig. 1
Four genes	B2 M	Coding	1438	GTR + G	6	Gamma	B
	HBB	Coding	1224	GTR + G	6	Gamma	B
	IRBP	Coding	1839	HKY + G	2	Gamma	A
	vWF	Coding	891	HKY + G	2	Gamma	B
Wildman et al.	Xq221	Non-coding	666	GTR + G	6	Gamma	A
	1p311	Non-coding	627	GTR	6	Equal	A
	10q231	Non-coding	666	HKY	2	Equal	A
	8q231	Non-coding	675	GTR + G	6	Gamma	B
	6p223	Non-coding	672	GTR + G	6	Gamma	B
	2p223	Non-coding	1149	HKY + I + G	2	Gamma	A
	10p1233	Non-coding	513	HKY	2	Equal	A
	3q222	Non-coding	560	HKY + G	2	Gamma	C
	3p13	Non-coding	678	GTR + G	6	Gamma	A
	1q313	Non-coding	693	GTR + G	6	Gamma	C
Perelman et al.	ABCA1_INTRON	Non-coding	560	GTR + G	6	Gamma	A
	AFF2_INTRON	Non-coding	500	GTR + G	6	Equal	A
	AXIN1_EXON	Coding	854	GTR + I + G	6	Gamma	B
	BRCA2_EXON	Coding	1252	GTR + G	6	Gamma	B
	CNR1_EXON	Coding	997	GTR + I + G	6	Gamma	B
	DMRT1_INTRON	Non-coding	537	GTR	6	Equal	A
	EDG1_EXON	Coding	967	HKY + I + G	2	Gamma	C
	NEGR1_INTRON	Non-coding	538	GTR + G	6	Gamma	A
	NPAS3_INTRON	Non-coding	605	GTR + G	6	Gamma	A
	NPAS3.2_INTRON	Non-coding	650	GTR + G	6	Gamma	A
	RAG1_EXON	Coding	1071	GTR + I + G	6	Gamma	C
	RPGRIP1_EXON1	Coding	431	HKY + G	2	Equal	B
	SGMS1_UTR	Non-coding	463	HKY + G	2	Gamma	A
	SIM1_INTRON	Non-coding	646	GTR + G	6	Gamma	A
	TYR_EXON	Coding	475	HKY + I	2	Equal	A
	USH2A_INTRON	Non-coding	605	GTR + G	6	Gamma	A
	ZFX_INTRON	Non-coding	811	GTR + G	6	Gamma	A
	ZIC3_UTR	Non-coding	433	GTR	6	Equal	A
	BCHE_EXON	Coding	984	GTR + G	6	Gamma	B
	DCTN2_UTR	Non-coding	528	GTR + I + G	6	Gamma	A
	POLA1_INTRON	Non-coding	604	GTR + G	6	Gamma	B
	RAB6IP1_INTRON	Non-coding	557	GTR + G	6	Gamma	B

substitution rates to vary across branches according to an uncorrelated lognormal distribution (Drummond et al., 2006). Because the fossil record for platyrrhines is still widely debated, we set the mean of branch rates to 1 (ucld.mean); therefore, time was measured in arbitrary units, providing a relative measure of divergence times. We then computed the maximum credibility tree in TreeAnnotator 1.4.8 (Drummond and Rambaut, 2007).

FigTree v1.3.1 was used to plot all of the gene and species trees.

3. Results

The species tree for the combined dataset supports (with high posterior probabilities) the previous division of the platyrrhines into five clades (Atelidae, Pitheciidae, Callitrichinae and Cebinae plus Aotinae [the Cebidae family]) and corroborates a closer phylogenetic relationship between Atelidae and Cebidae (Fig. 1; Opazo et al., 2006; Wildman et al., 2009). The topology of the trees obtained via concatenated and multispecies coalescent Bayesian analyses are identical. However, the multispecies coalescent Bayesian analyses of the combined dataset including missing loci do not converge after 200,000,000–600,000,000 generations (results not shown). Within the Atelidae, *Alouatta* is a sister group of a clade that includes *Ateles*, *Brachyteles* and *Lagothrix*, in which *Brachyteles* and *Lagothrix* are clustered together. Within the Cebidae family, we find the Cebinae branch, which includes the *Cebus* and *Saimiri* genera, as well as the Callitrichinae branch, which is formed by the *Saguinus*, *Leontopithecus*, *Callithrix* and *Callimico* genera. *Aotus* is a branch external to the Cebidae family. Within Pitheciidae, *Cacajao* and *Chiropotes* are clustered together, whereas *Pithecia* and *Callicebus* are not. *Pithecia* is a sister group of the former clade, and *Callicebus* is the most external clade (Fig. 1).

In the species tree of the first sub-dataset, which includes four nuclear genes, most of the clades are strongly supported with posterior probabilities higher than 0.95. Therefore, there are some relationships characterized by lower posterior probabilities (i.e., Cebinae and *Aotus*; Fig. 2A). The tree is mainly concordant with our species tree based on the combined dataset; however, in the species tree based on the four sequences *Aotus* is a related to Cebinae (Fig. 2A; Wildman et al., 2009; Opazo et al., 2006). The lower support and topological difference in the relationship between *Aotus* and the Cebinae branch is due to the trees of the four genes being characterized by differences in their topological relationships. Whereas the *IRBP* tree shows a phylogenetic relationship between *Aotus* and the Cebidae family (Supplementary Fig. 1A), the *vWF*, *B2M* and *HBB* trees display a closer relationship between *Aotus* and the Cebinae clade (Supplementary Fig. 1B). The concatenated BEAST analyses show a topology similar to the ∗BEAST species tree estimation, but all of the relationships exhibit posterior probabilities higher than 0.95 (Fig. 2B).

The ∗BEAST analysis of the second sub-dataset, containing 10 non-coding sequences (Wildman et al., 2009) yields a tree in which most of the clades are strongly supported with high posterior probabilities (>0.95), but in which some relationships are characterized by lower posterior probabilities (i.e., Cebinae and Callitrichinae; Atelidae and Cebidae; Fig. 3A). The topology of this tree is concordant with our species tree based on the combined dataset, with *Aotus* as a branch external to the Cebidae family. The gene trees of the 10 non-coding sequences are also characterized by differences in

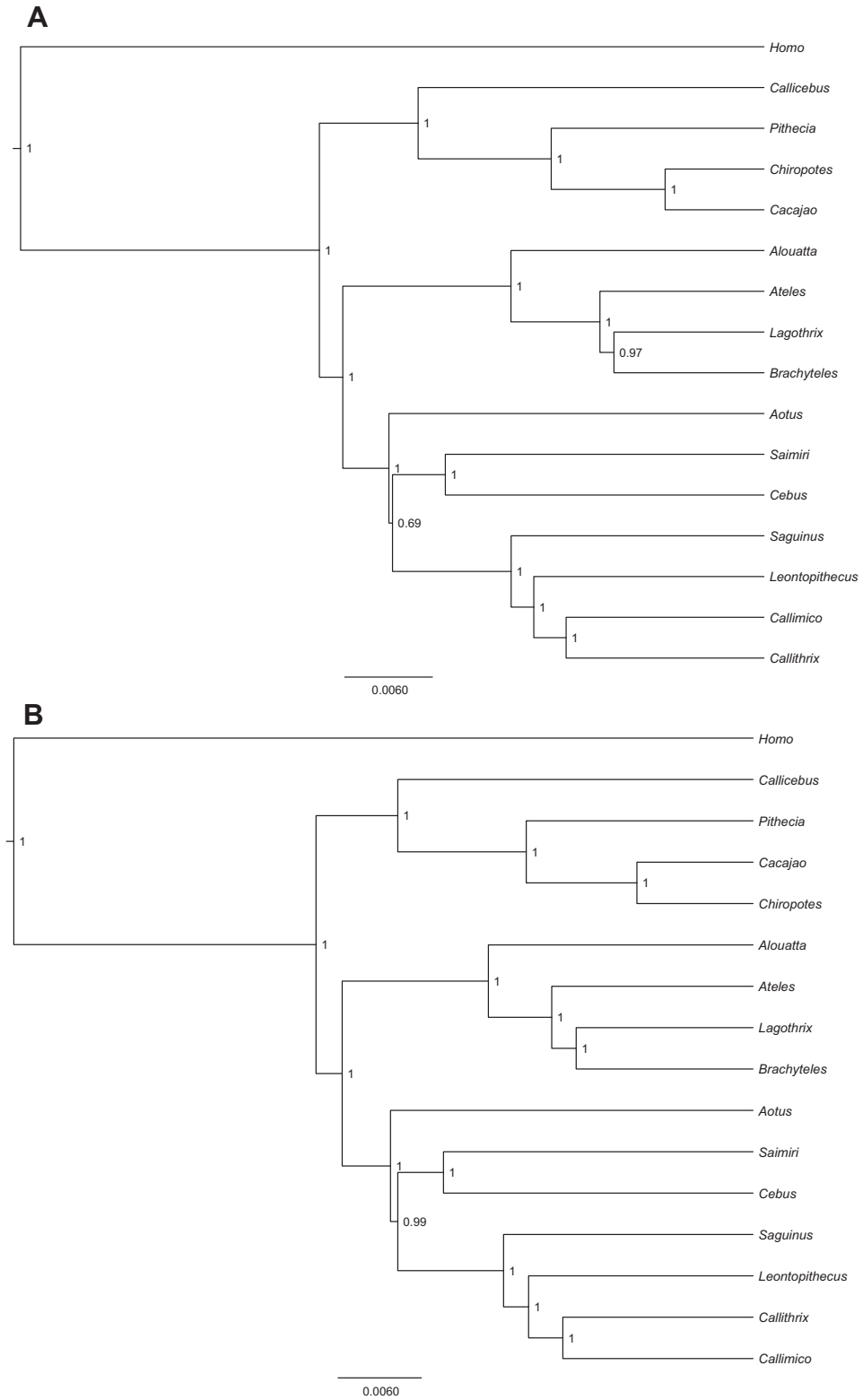

Fig. 1. Phylogenetic trees based on the combined dataset (68 loci and 47,233 bp or 36 loci and 27,360 bp) for 15 genera. (A) *BEAST species tree. (B) BEAST concatenated tree. Node posterior probabilities and branch lengths are indicated on the graphs.

their topological relationships. The tree for the 1p31.1, 2p22.3, 3p13, 10p12.33, 10q23.1 and Xq22.1 sequences shows that *Aotus* is related to the Cebidae family (Supplementary Fig. 1A), whereas the tree for the 6p22.3 and 8q23.1 sequences shows that *Aotus* is related to the Cebinae (Supplementary Fig. 1B). In addition, the tree for the 1q31.3 and 3q22.2 sequences shows that *Aotus* is related to the Callitrichinae (Supplementary Fig. 1C). The 8q23.1, 10q23.1 and 3p13 sequences also differ in the phylogenetic relationship

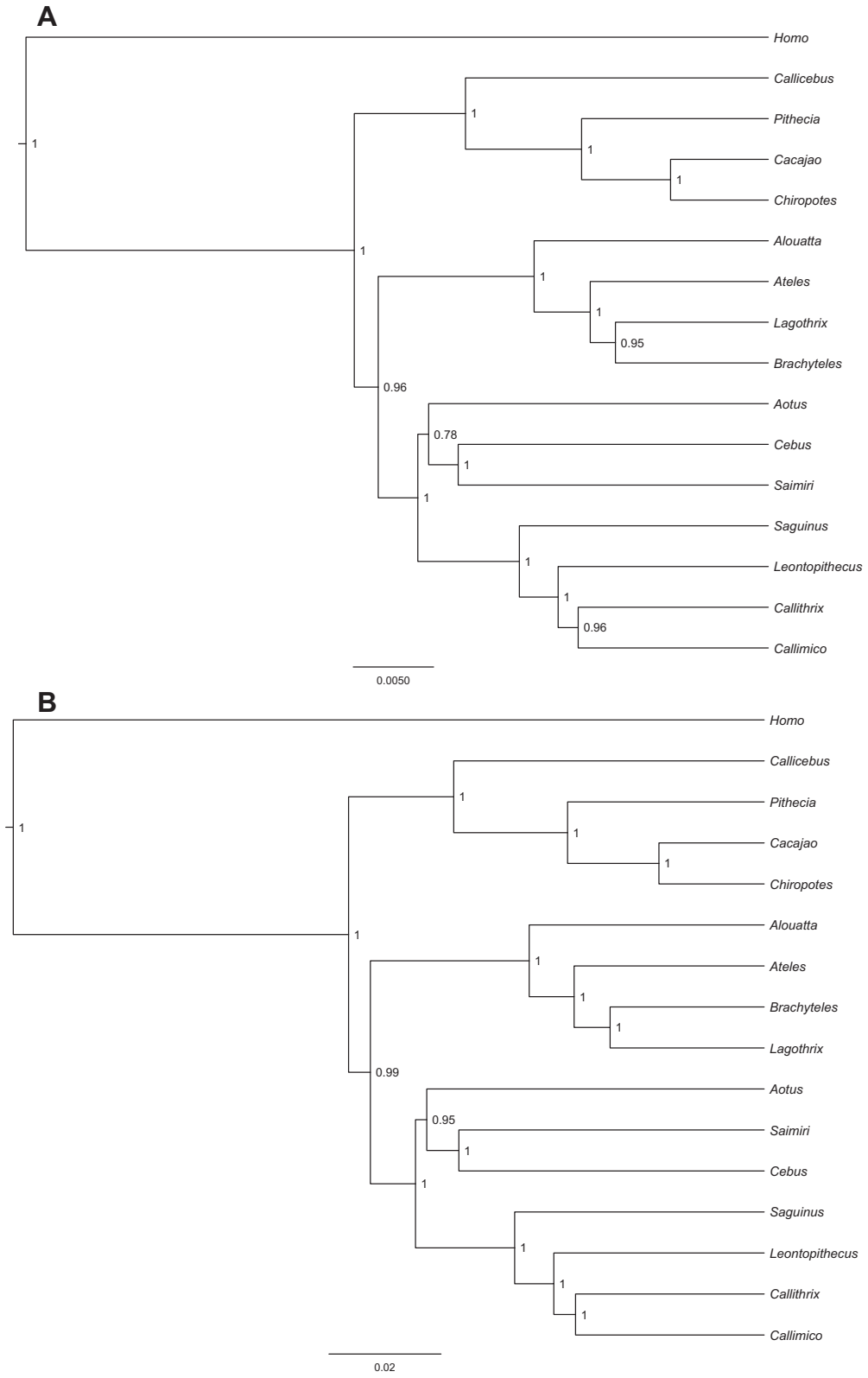

Fig. 2. Phylogenetic trees based on the sub-dataset of four coding sequences. (A) ∗BEAST species tree. (B) BEAST concatenated tree. Node posterior probabilities and branch lengths are indicated on the graphs.

they show among the platyrrhine families. The concatenated BEAST analyses of the 10 non-coding sequences present a topology similar to the ∗BEAST species tree estimation, but these relationships have posterior probabilities higher than those for the ∗BEAST species tree (Fig. 3B).

Finally, the BEAST analysis of the third sub-dataset, including 54 sequences (Perelman et al., 2011), yields a tree in which most of the clades are strongly supported with high posterior probabilities and where the topology is mainly concordant with our species tree based on the combined dataset. The difference between these two

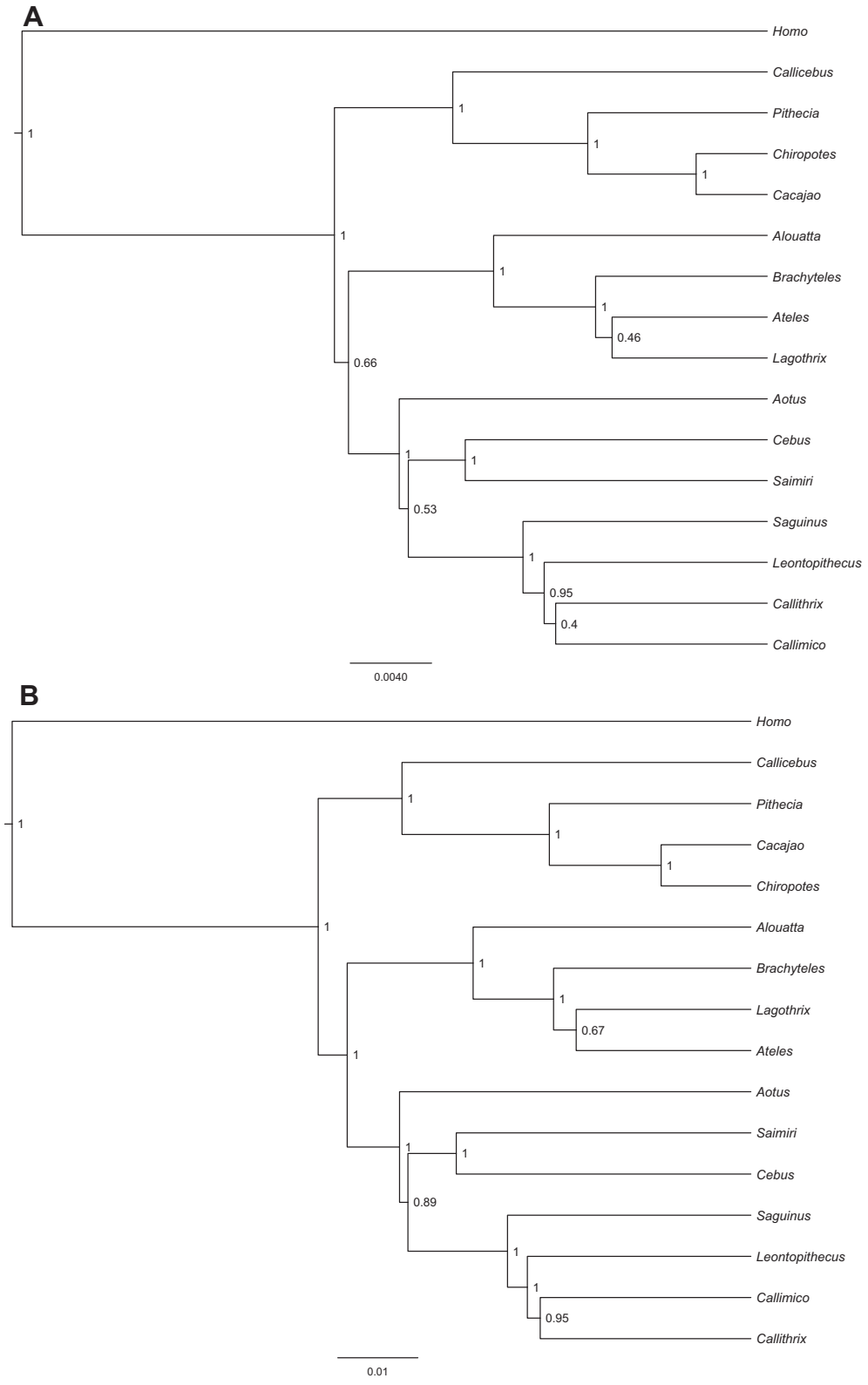

Fig. 3. Phylogenetic trees based on the sub-dataset of ten non-coding sequences (Wildman et al., 2009). (A) *BEAST species tree. (B) BEAST concatenated tree. Node posterior probabilities and branch lengths are indicated on the graphs.

trees is that in the first tree, *Aotus* is a branch external to the Cebidae family, whereas in the second, this genus constitutes a branch related to Callitrichinae (Supplementary Figs. 2A and 2B). Conversely, the *BEAST analysis of the third sub-dataset without missing data, including 22 coding and non-coding sequences (Perelman et al., 2011), yields a tree in which most of the clades are strongly supported with high posterior probabilities, but with one relationship being characterized by a very low posterior probability (i.e.,

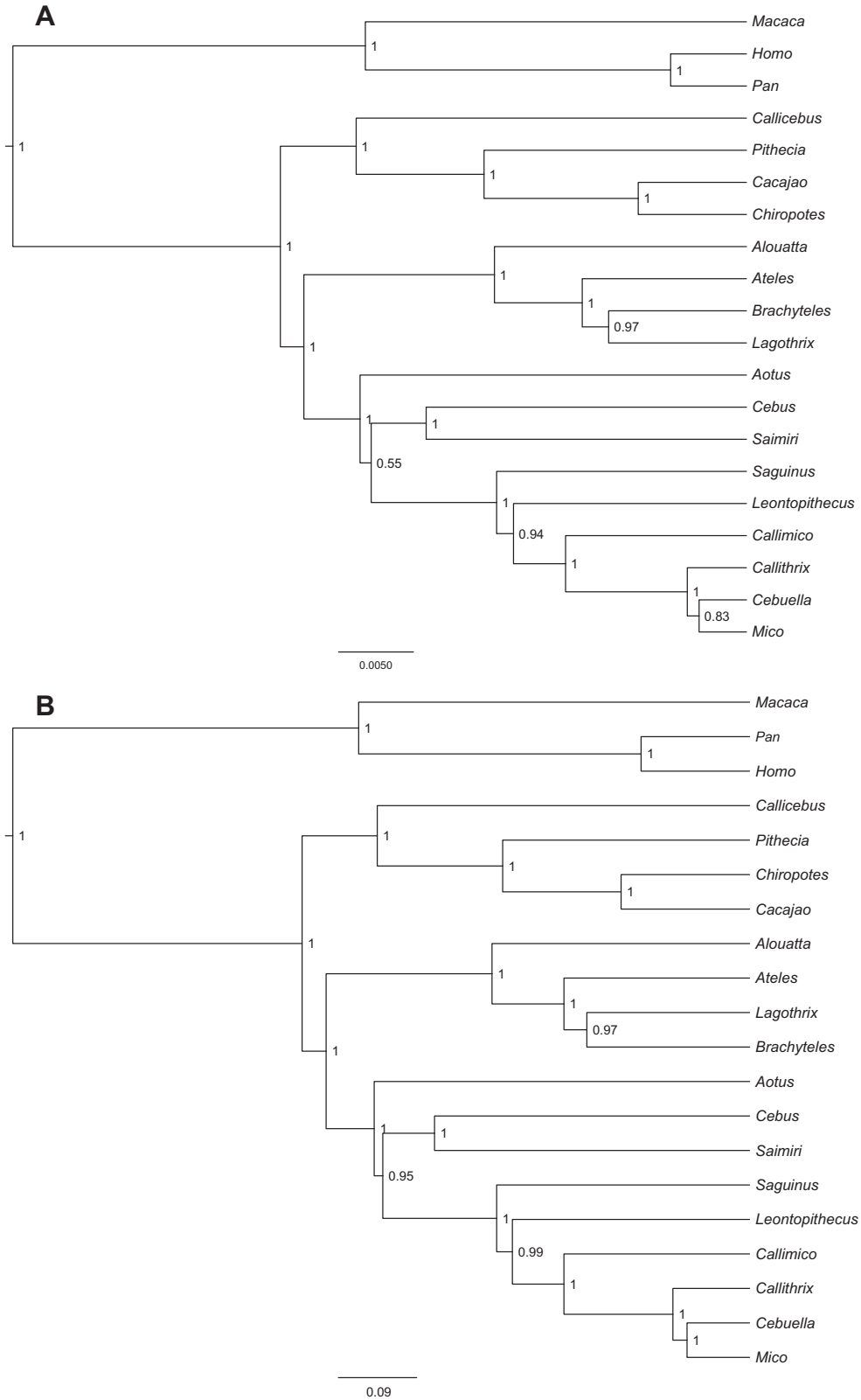

Fig. 4. Phylogenetic trees based on the sub-dataset of 22 coding and non-coding sequences (Perelman et al., 2011). (A) *BEAST species tree. (B) BEAST concatenated tree. Node posterior probabilities and branch lengths are indicated on the graphs.

between Cebinae and Callitrichinae; Fig. 4A). Moreover, this analysis is concordant with the species tree based on the combined and Wildman dataset. The gene trees of the 22 coding and non-coding sequences are also characterized by differences in their topological relationships. The tree for the sequences ABCA1_INTRON, AFF2_INTRON, DCTN2_UTR, DMRT1_INTRON, NEGR1_INTRON, NPAS3.2_INTRON, NPAS3_INTRON, SGMS1_UTR, SIM1_INTRON, TYR_EXON, USH2A_INTRON, ZFX_INTRON and ZIC3_UTR show that *Aotus* is related to the Cebidae family (Supplementary Fig. 1A), whereas the tree for the sequences AXIN1_EXON, BCHE_EXON,

BRCA2_EXON, CNR1_EXON, POLA1_INTRON, RAB6IP1_INTRON and RPGRIP1_EXON1 show that *Aotus* is related to the Cebinae (Supplementary Fig. 1B). Finally, the tree for the EDG1_EXON and RAG1_EXON sequences shows *Aotus* to be related to the Callitrichinae (Supplementary Fig. 1C). The concatenated BEAST analyses indicate a similar topology to the ∗BEAST species tree estimation but all of the relationships have posterior probabilities higher than 0.95 (Fig. 4B).

Because there are some differences in the implementation of the Bayesian MCMC method for the multispecies coalescent (see Heled and Drummond, 2010) and the partitioned Bayesian analysis of concatenated sequences (Drummond and Rambaut, 2007), we reanalyzed the three sub-datasets using BEST 2.3 (Liu, 2008) and MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) software. The results from both BEST 2.3 (Liu, 2008) and MrBayes 3.1.2 are generally similar to the BEAST and ∗BEAST trees (Supplementary Fig. 3). However, as discussed previously, ∗BEAST runs at a higher computational speed (Heled and Drummond, 2010). Similar results were also obtained using the maximum likelihood implemented in Mega 5.05 (Tamura et al., 2011; results not shown).

Moreover, as it has been noted, the tendency to root *Aotus* within the Cebidae in previous molecular studies could be caused by long-branch attraction (Rosenberger and Tejedor, in press); please note that the effects of long-branch attraction and taxon sampling on the Bayesian method of species tree estimation remain unexplored (Degnan and Rosenberg, 2009). We reanalyzed the three sub-datasets excluding the *Saimiri* and *Cebus* genera (the long-lived lineages of platyrrhines; Bergsten, 2005; Rosenberger and Tejedor, in press). The resulting concatenated and species trees are generally similar to the tree for the total dataset, placing *Aotus* as a sister clade of the Callitrichinae, but these relationships have posterior probabilities of 1 (Supplementary Fig. 4).

The topological relationships among the platyrrhines observed in previous molecular studies, particularly regarding the relationship of *Aotus* with the Cebidae clade, could be caused by an inadequate species representation (as noted by one reviewer). We reanalyzed the Perelman et al. sub-dataset using a concatenated BEAST approach for 42 (we only used the species with more than 25,000 bp available for the BEAST analysis) and 14 (we analyzed one species for each genera) platyrrhine species. The resulting concatenated trees are generally similar to each other and to the original tree of Perelman et al. (2011), placing *Aotus* as a sister clade of the Callitrichinae (Supplementary Figs. 2A and 2B). We also re-analyzed the 22 coding and non-coding sequences in the Perelman sub-dataset using 27 and 14 Platyrrhine species and 15,068 bp. The resulting concatenated trees are generally similar to each other, placing *Aotus* as a sister clade of the Cebidae (Supplementary Figs. 2C and 2D).

4. Discussion

Because of the putative problems generated when using a single gene sequence or a concatenation of a few gene sequences to estimate a species tree, in this study, we estimated the platyrrhine species tree for each dataset combining multiple gene trees (due to the number of gene sequences involved, the combined dataset generated our main and most robust phylogeny) and a multispecies coalescent model (Heled and Drummond, 2010). Our species trees obtained with ∗BEAST from the Wildman and summarized Perelman datasets show a topology similar to the species tree obtained with the combined dataset and to the majority of gene trees estimated for each dataset, indicating a closer phylogenetic relationships between Atelidae and Cebidae as well as between *Aotus* and the Cebidae clade (Figs. 1–4). However, the posterior probabilities are lower for the conflictive tree nodes. This result could be explained by the fact that some gene trees showed no concordant topologies or gene histories (Supplementary Fig. 1; Liu et al., 2008). Similar results have been observed previously in species tree estimations for other taxa, suggesting that concatenation may present identical topologies to species tree estimation based on the multispecies coalescent model but overestimate the posterior probabilities (e.g., Liu et al., 2008; Belfiore et al., 2008; Edwards, 2009; Clemente-Carvalho et al., 2011). Estimating support for gene tree or concatenated tree branches is different than estimating support for species tree branches, as the last process is principally related to the number of genes presenting a particular relationship in a tree node (Liu et al., 2008). It is important to note that when we eliminate the *Saimiri* and *Cebus* genera (the longest branch after *Aotus* in the platyrrhine tree) from the datasets, the resulting species tree places *Aotus* as a sister clade of the Callitrichinae and shows a closer phylogenetic relationships between Atelidae and Cebidae; however, these relationships have posterior probabilities of 1 and 0.98, respectively. It is also of note that incomplete taxon sampling (i.e., using only one species per genus) was apparently not a problem in our species tree estimations (Supplementary Fig. 2); but missing data could generate problems for species tree estimation using the multispecies coalescent model (see Liu et al., 2008; Edwards, 2009).

Our gene tree estimations for the platyrrhines show that the studied genes display different topologies, probably explaining the difference in tree topology generated with the datasets that were used in previous platyrrhine analyses. The most frequent pattern is a gene tree that shows a closer phylogenetic relationship between Atelidae and Cebidae and a closer relationship between *Aotus* and the Cebidae clade (Supplementary Fig. 1A). The existence of a discrepancy between gene trees and the species tree has been known since the 1980s (Maddison, 1997; Felsenstein, 2004; Rannala and Yang, 2008; Knowles and Kubatko, 2010). These differences can be related to processes such as incomplete lineage sorting, horizontal gene transfer and selection, among others (Maddison, 1997; Degnan and Rosenberg, 2009; Knowles and Kubatko, 2010). Among these processes, incomplete lineage sorting is the most widely observed, and is this process assumed in ∗BEAST and BEST (Liu, 2008; Liu et al., 2008; Heled and Drummond, 2010). Although incomplete lineage sorting is common in species trees with low divergence (Belfiore et al., 2008; Clemente-Carvalho et al., 2011), it can also occur in species trees with deep divergence. This phenomenon could be particularly frequent in ancient rapid radiations (Degnan and Rosenberg, 2009), in which short ancient tree branches can be common, such as in the platyrrhine radiation. Our Bayesian platyrrhine species trees also suggest this scenario, as there is a long separation between the divergence of *Aotus* and the divergence of most of the platyrrhine genera along, with a short interval between the separation of *Aotus* and the Cebidae crown (Figs. 1–4).

The concatenated BEAST analyses show a similar topology with high posterior probabilities (higher than 0.95) for the majority of gene trees estimated for each dataset and the species tree estimation; therefore, the concatenated estimation based on the four-gene dataset differs from the combined, Wildman and summarized Perelman datasets regarding the position of *Aotus* (Figs. 1B, 2B, 3B and 4B). It is common practice in phylogenetic analyses of platyrrhines and other animal groups to use a single gene sequence or a concatenation of a few gene sequences to estimate a phylogenetic tree (Degnan and Rosenberg, 2009; Lemey et al., 2009; Heled and Drummond, 2010). This practice could be particularly problematic for species tree estimation because while the estimation of gene trees is based on DNA sequence data, a species tree is the unobserved tree of genealogical relationships among species. Therefore, the estimation of a species tree is based on the gene trees (Rannala and Yang, 2008). It was suggested that using concat-

enated DNA sequence data, we could generate trees that display, or are dominated by the topology of the gene or genes with the longest sequences. Moreover, although the use of a gene or a concatenation of gene sequences assumes that the data have evolved under different mutation rates and models for different sites, they also assume that the gene sequences have evolved according to a single evolutionary tree (Degnan and Rosenberg, 2009). When the genes have different evolutionary histories, the last assumption is violated, and theoretical studies suggest that such a violation can result in a well-supported but sometimes incorrect species tree (Degnan and Rosenberg, 2009; Heled and Drummond, 2010). Our results indicate that the previous differences between previous phylogenetic trees proposed for the platyrrhines are likely to be attributable to the particular gene tree or concatenated gene tree used in a given study. In particular, the differences between our phylogenetic tree based on the dataset of four nuclear genes and each of the other concatenation analyses included in the present report (Figs. 1–4) may be related to the fact that we used fewer genes in the first dataset, among which three genes are DNA coding genes with longer sequences (Table 1), but with an infrequent topology in platyrrhines (Supplementary Fig. 1; see Clemente-Carvalho et al., 2011). However, our species tree based on the multispecies coalescent model exhibits the same topology as the concatenated analyses, showing that the species tree topology estimation is also dependent on the studied gene trees. As noted above, the two phylogenetic approaches compared in this study differ in the statistical support obtained for conflictive nodes, which is lower under the multispecies coalescent approach than in the concatenated analyses, in accordance with the idea that the former approach would better incorporate gene tree variation into the phylogenetic analysis (Degnan and Rosenberg, 2009; Edwards, 2009).

Our species trees based on many sequences (i.e., the combined, summarized Perelman and Wildman datasets) are similar to those produced in some recent concatenated analyses. In particular, the topology is virtually identical to the concatenated tree displayed in Fig. 2 of Wildman et al. (2009) and shows several genera with relatively long branches, as noted previously (Rosenberger, 2002; Rosenberger et al., 2009). However, the topology of our species tree differs from the results of Opazo et al. (2006) and Perelman et al. (2011) regarding the placement of *Aotus*. Our Bayesian species tree based on the Perelman and Wildman datasets also differs from the morphological tree reported by Rosenberger (1992) in the placement of *Aotus* and in the relationships between Pitheciidae and the other clades. Moreover, our species tree differs greatly from other morphological trees (e.g., Ford, 1986; Kay, 1990). The discrepancy between morphological and molecular data and species trees has been studied for decades and is especially problematic for closely related taxa (Felsenstein, 1985; Heled and Drummond, 2010) and/or taxa under adaptive radiation (Gavrilets and Losos, 2009; Losos and Mahler, 2010). The phylogenetic history of one morphological character or a cluster of morphological characters may differ from the species tree because of processes such as selection, among others. Moreover, analogous to the gene tree and concatenated analyses of sequences, the topology of concatenated morphological phylogenies could be determined by the characters that dominate the matrix (Rosenberger, 2002).

In summary, we used a Bayesian method for multispecies coalescent modeling to perform joint estimation of a species tree topology, relative divergence times and gene trees from multiple sequences obtained for 15–17 platyrrhine genera. We generated a robust estimation of the phylogenetic tree of platyrrhine genera based on a multispecies coalescent model and concatenated method using all of the available datasets. The topological differences observed between our species tree and the previous phylogenetic trees are likely related to the use of particular gene sequences. Thus, from a methodological point of view, our gene and species tree results support the increasingly accepted idea that multiple independent sequences from a few individuals and a Bayesian multispecies coalescent framework can generate species tree estimations that are more realistic—in terms of their statistical support—than just one or a few molecular loci from a large number of taxa and a concatenated framework (Rannala and Yang, 2008; Heled and Drummond, 2010). Moreover, this new approach suggests that when studying deep phylogenetic divergences, it is important to maximize the sampling effort not only for the number of base pairs, but mainly, for the number of independently segregating genes (e.g., the Wildman dataset; Edwards, 2009). From a more conceptual point of view, the observed discrepancy among gene trees, morphological trees and species trees could be related to several different evolutionary and ecological processes, including selection, plasticity, horizontal transfer, lineage sorting and gene duplication/extinction (Felsenstein, 1985; Maddison, 1997; Rannala and Yang, 2008). Many of the sources of inconsistency between the histories described based on morphological or molecular phenotypic traits and the history of species divergence (i.e., the species trees) could be a main focus of further research on platyrrhines and other mammalian groups.

Acknowledgments

We thank M. Tejedor for providing helpful critical comments on an early version of the complete manuscript of this work. We also thank editors from American Journal Experts for help with the English version of the manuscript. We are deeply indebted to two anonymous reviewers who contributed greatly to improving the study. This research is supported by Grants from the Fundo de Apoio ao Ensino, a Pesquisa e a Extensão (FAEPEX – UNICAMP). S.I. Perez is supported by a research fellowship from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and FONCyT (PICT-2011-0307). Julia Klaczko is supported by a postdoctoral fellowship from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – 2011/17070-6). S.F. dos Reis is partially supported by grants from the Fundação de Amparo a Pesquisa do Estado de São Paulo and a research fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.07.014.

References

Belfiore, N.M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: Echimyidae). Syst. Biol. 57, 294–310.
Bergsten, J., 2005. A review of long-branch attraction.. Cladistics 21, 163–193.
Clemente-Carvalho, R., Klaczko, J., Perez, S.I., Alves, A., Haddad, C., dos Reis, S.F., 2011. Molecular phylogenetic relationships and phenotypic diversity in miniaturized toadlets, genus *Brachycephalus* (Amphibia: Anura: Brachycephalidae). Mol. Phylogenet. Evol. 61, 79–89.
Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.
Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? Evolution 63, 1–19.
Felsenstein, J., 1985. Phylogenies and the comparative method. Am. Nat. 125, 1–15.
Felsenstein, J., 2004. Inferring Phylogenies. Sinauer.
Fleagle, J.G., 1999. Primate Adaptation and Evolution, second ed. Academic Press, New York.
Ford, S.M., 1986. Systematics of the New World monkeys. In: Swindler, D.R., Erwin, J. (Eds.), Comparative Primate Biology, Systematics, Evolution and Anatomy, vol. 1. Alan R. Liss, New York, pp. 73–135.

Freckleton, R.P., Cooper, N., Jetz, W., 2011. Comparative methods as a statistical fix: the dangers of ignoring an evolutionary model. Am. Nat. 178, E10–E17.
Gavrilets, S., Losos, J.B., 2009. Adaptive radiation: contrasting theory with data. Science 323, 732–737.
Goodman, M., Porter, C.A., Czelusniak, J., Page, S.L., Schneider, H., Shoshani, J., Gunnell, G., Groves, C.P., 1998. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. Mol. Phylogenet. Evol. 9, 585–598.
Hall, T., 2004. BioEdit 7.0.0. <http://www.mbio.ncsu.edu/bioedit/bioedit.html>.
Heled, J., Drummond, A.J., 2010. Bayesian inference of Species Trees from multilocus data. Mol. Biol. Evol. 27, 570–580.
Heled, J., Drummond, A.J., Xie, W., 2011. Bayesian inference of Species Trees from multilocus data using ∗BEAST. <http://www.beast.bio.ed.ac.uk/Tutorials>.
Horovitz, I., Zardoya, R., Meyer, A., 1998. Platyrrhine systematics: a simultaneous analysis of molecular and morphological data. Am. J. Phys. Anthropol. 106, 261–281.
Kay, R.F., 1990. The phyletic relationships of extant and fossil Pitheciinae. J. Hum. Evol. 19, 175–208.
Knowles, L.L., Kubatko, L.S., 2010. Estimating Species Trees: Practical and Theoretical. Wiley-Blackwell, New York.
Leaché, A.D., Rannala, B., 2011. The accuracy of species tree estimation under simulation: A comparison of methods. Syst. Biol. 60, 126–137.
Lemey, M., Salemi, M., Vandamme, A.-M., 2009. The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge University Press, Cambridge.
Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24, 2542–2543.
Liu, L., Pearl, D.K., Brumfield, R.T., Edwards, S.V., 2008. Estimating species trees using multiple-allele DNA sequence data. Evolution 62, 2080–2091.
Losos, J.B., 2011. The limitations of phylogenies in comparative biology. Am. Nat. 177, 709–727.
Losos, J.B., Mahler, D.L., 2010. Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. In: Bell, M.A., Futuyma, D.J., Eanes, W.F., Levinton, J.S. (Eds.), Evolution Since Darwin: The First 150 Years. Sinauer Associates, Sunderland, MA, pp. 381–420.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.
Norconk, M.A., Wright, B.W., Conklin-Brittain, N.L., Vinyard, C.J., 2009. Mechanical and nutritional properties of food as factors in Platyrrhine dietary adaptations. In: Garber, P.A., Estrada, A., Bicca-Marques, J.C., Heymann, E.W., Strier, K.B. (Eds.), South American Primates, Developments in Primatology: Progress and Prospects. Springer-Verlag, New York, pp. 279–319.
Opazo, J.C., Wildman, D.E., Prychitko, T., Johnson, R.M., Goodman, M., 2006. Phylogenetic relationships and divergence times among New World monkeys (Platyrrhini, Primates). Mol. Phylogenet. Evol. 40, 274–280.
Perelman, P., Johnson, W.E., Roos, C., Seuánez, H.N., Horvath, J.E., Moreira, M.A.M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M.P.C., Silva, A., O'Brien, S.J., Pecon-Slattery, J., 2011. A molecular phylogeny of living Primates. PLoS Genet. 7, e1001342.
Perez, S.I., Klaczko, J., Rocatti, G., dos Reis, S.F., 2011. Patterns of cranial shape diversification during the phylogenetic branching process of New World monkeys (Primates: Platyrrhini). Evol. Biol. 24, 1826–1835.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4 software. <http://www.beast.bio.ed.ac.uk/Tracer>.
Rannala, B., Yang, Z., 2008. Phylogenetic inference using whole genomes. Annu. Rev. Genom. Hum. Genet. 9, 217–231.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Rosenberger, A.L., 1984. Fossil New World Monkeys dispute the molecular clock. J. Hum. Evol. 13, 737–742.
Rosenberger, A.L., 1992. The evolution of feeding niches in New World Monkeys. Am. J. Phys. Anthropol. 88, 525–562.
Rosenberger, A.L., 2002. Platyrrhine paleontology and systematics: the paradigm shifts. In: Hartwig, W.C. (Ed.), The Primate Fossil Record. Cambridge University Press, Cambridge, pp. 151–160.
Rosenberger, A.L., Tejedor, M.F., Cooke, S., Halenar, L., Pekkar, S., 2009. Platyrrhine ecophylogenetics, past and present. In: Garber, P.A., Estrada, A., Bicca-Marques, J.C., Heymann, E.W., Strier, K.B. (Eds.), South American Primates: Comparative Perspectives in the Study of Behavior. Ecology and Conservation, Springer, New York, pp. 69–113.
Rosenberger, A.L., Tejedor, M.F, in press. The misbegotten: long lineages, long branches and the interrelationships of *Aotus*, *Callicebus* and the saki-uakaris. In: Barnett, A.L., Veiga, S., Ferrari, S., Norconk, M.N. (Eds.), Evolutionary Biology and Conservation of Titis, Sakis and Uacaris. Cambridge University Press, Cambridge
Schneider, H., 2000. The current status of the NewWorld Monkey phylogeny. An. Acad. Bras. Ci. 72 (2).
Schneider, H., Canavez, F.C., Sampaio, I., Moreira, M.Â.M., Tagliaro, C.H., Seuánez, H.N., 2001. Can molecular data place each neotropical monkey in its own branch? Chromosoma 109, 515–523.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Tejedor, M.F., 2008. The origin and evolution of neotropical primates. Arq. Mus. Nac. 66, 251–269.
Wiens, J.J., 2009. Estimating rates and patterns of morphological evolution from phylogenies: lessons in limb lability from Australian Lerista lizards. J. Biol. 8, 19.
Wildman, D.E., Jameson, N.M., Opazo, J.C., Yi, S.V., 2009. A fully resolved genus level phylogeny of neotropical primates (Platyrrhini). Mol. Phylogenet. Evol. 53, 694–702.
Wilkinson, R.D., Steiper, M.E., Soligo, C., Martin, R.D., Yang, Z., Tavaré, S., 2011. Dating primate divergences through an integrated analysis of palaeontological and molecular data. Syst. Biol. 60, 16–31.