Phylogeny and historical biogeography of gnateaters (Passeriformes, Conopophagidae) in the South America forests

Molecular Phylogenetics and Evolution, v.79, p.422-432, 2014
http://www.producao.usp.br/handle/BDPI/46195

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo
Phylogeny and historical biogeography of gnat-eaters (Passeriformes, Conopophagidae) in the South America forests

Henrique Batalha-Filho a,b,*, Rodrigo O. Pessoa c, Pierre-Henri Fabre d,e, Jon Fjeldså d, Martin Irestedt f, Per G.P. Ericson f, Luís F. Silveira g, Cristina Y. Miyaki a

a Departamento de Genética e Biologia Evolutiva, Instituto de Biocências, Universidade de São Paulo, São Paulo, SP, Brazil
b Centro de Ciências Biológicas e da Saúde, Universidade Estadual de Montes Claros, Montes Claros, MG, Brazil
c Museu de Zoologia, Universidade de São Paulo, São Paulo, SP, Brazil
d Center for Macroecology, Evolution and Climate at the Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
e Harvard Museum of Comparative Zoology, 26 Oxford Street, Cambridge, MA 02138, USA
g Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, SE-104 05 Stockholm, Sweden

Abstract

We inferred the phylogenetic relationships, divergence time and biogeography of Conopophagidae (gnat-eaters) based on sequence data of mitochondrial genes (ND2, ND3 and cytb) and nuclear introns (TGFβ2 and G3PDH) from 45 tissue samples (43 Conopophaga and 2 Pittasoma) representing all currently recognized species of the family and the majority of subspecies. Phylogenetic relationships were estimated by maximum likelihood and Bayesian inference. Divergence time estimates were obtained based on a Bayesian relaxed clock model. These chronograms were used to calculate diversification rates and reconstruct ancestral areas of the genus Conopophaga. The phylogenetic analyses support the reciprocal monophyly of the two genera, Conopophaga and Pittasoma. All species were monophyletic with the exception of C. lineata, as C. lineata cearae did not cluster with the other two C. lineata subspecies. Divergence time estimates for Conopophagidae suggested that diversification took place during the Neogene, and that the diversification rate within Conopophaga clade was highest in the late Miocene, followed by a slower diversification rate, suggesting a diversity-dependent pattern. Our analyses of the diversification of family Conopophagidae provided a scenario for evolution in Terra Firme forest across tropical South America.

The spatio-temporal pattern suggests that Conopophaga originated in the Brazilian Shield and that a complex sequence of events possibly related to the Andean uplift and infilling of former sedimentation basins and erosion cycles shaped the current distribution and diversity of this genus.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

The origin of the South American rain forest biota has fascinated naturalists since Wallace’s days (Wallace, 1852), and several hypotheses have been evoked to explain the history of diversification (review in Moritz et al., 2000) in this highly diverse region (Myers et al., 2000). For many years the Pleistocene refuge hypothesis (Haffer, 1969; Vanzolini and Williams, 1970) was assumed to represent the principal mechanism to explain diversification in the Amazon basin and other lowland rainforest regions in South America (Carna val and Moritz, 2008; Whitmore and Prance, 1987). However, not all paleoecological data supported the

* Corresponding author. Address: Departamento de Zoologia, Instituto de Biologia, Universidade Federal da Bahia, Rua Barão de Geronemooabo, 147, Ondina, 40170-290 Salvador, BA, Brazil. Fax: +55 (71) 32836511.
E-mail address: henrique.batalha@outlook.com (H. Batalha-Filho).

http://dx.doi.org/10.1016/j.ympev.2014.06.025
1055-7903/ © 2014 Elsevier Inc. All rights reserved.
Weir and Price, 2011). However, lineage diversification events seem to have occurred during all the Tertiary and Quaternary, even including the Pleistocene (Rull, 2008, 2011a, 2011b).

The hypothesis by Wallace (1852) that rivers could be barriers to dispersal and that this could lead to speciation, received new attention as geological evidence suggested a recent (Plio-Pleistocene) establishment of the Amazonian drainage channel to the Atlantic Ocean (Campbell et al., 2006; Latrubesse et al., 2010). This scenario has been shown to be congruent with recent divergence of avian populations across the river (Fernandes et al., 2012; Maldonado-Coelho et al., 2013; Ribas et al., 2012). According to this scenario, the Purus Arch (at 62°W) formed an eastern limit bounding the Solimões wetland system in western Amazonia until 9.5–6 Myr (Latrubesse et al., 2010), when the Amazon River broke through this barrier to drain toward the Atlantic Ocean.

To further investigate these historical scenarios, the gnateaters (Passeriformes: Conopophagidae) can be a good study case. These birds are an old lineage of New World suboscines (~41 Ma; Ohlson et al., 2013) that, in turn, evolved within South America, with only a few subclades colonizing Central and North America as part of the great American interchange (Smith and Klicka, 2010; Weir et al., 2009). The gnateaters presumably form a monophyletic genus (Ohlson et al., 2013). They are endemic to the Neotropical region and associated with mesic Terra Firme forest. Therefore, they were likely limited by the distribution of flat and hydrologically unstable depositional landscapes of the Amazonian floodplains. The family was long defined to comprise only the genus Conopophaga (Whitney, 2003), but recent molecular phylogenies have placed Pittasoma (formerly Formicariidae) as sister to this genus (Moyle et al., 2009; Ohlson et al., 2013; Rice, 2005). Thus, Conopophagidae currently comprises two genera and ten species (Remsen et al., 2013) that live in the understory of humid forests (Ridgely and Tudor, 1994). The genus Pittasoma has two species (P. michleri and P. rufopileatum) distributed in Central America and northwestern South America (Chocó region). The genus Conopophaga includes eight species (C. lineata, C. aurita, C. roberti, C. peruviana, C. ardesiaca, C. castaneiceps, C. melanops, and C. melanogaster), mostly polytypic and all endemic to South America (Fig. 1). Given the distributions of gnateater species through most of South America’s tropical forests (Whitney, 2003), and the relative old age of the clade (Ohlson et al., 2013), we expect that the evolutionary history of this group should reflect the major events of landscape history in these regions. Thus, a reconstruction of the phylogenetic history of gnateaters would provide a useful framework to further test the origin of extant Neotropical biodiversity, as they evolved in “splendid isolation” within South America during the Tertiary (Ricklefs, 2002).

In this study, we infer the phylogenetic relationships, divergence time and biogeography of Conopophagidae by using mitochondrial and nuclear genes of a comprehensive taxon sampling of the family. In addition, we examine the diversification within Conopophaga by using a wide geographic sampling, and analyze its diversification rates through time, as well as its ancestral distribution. To shed light on the gnateaters’ evolution we want to answer two main questions: (i) what are the phylogenetic relationships among species and subspecies of Conopophaga, and how these relationships correspond to currently defined species limits? and (ii) can the tempo and sequence of speciation events be associated with described events in the landscape history of South America?

### 2. Materials and methods

#### 2.1. Taxon sampling and molecular methods

In order to recover the phylogenetic relationships within Conopophagidae, we analyzed 45 samples (43 Conopophaga and 2 Pittasoma) from the ingroup (Fig. 1; Table S1 in supplementary material), representing all currently recognized species of the family. For most species we included at least two individuals and for polytypic species most subspecies were sampled. In total we included 14 of the 18 recognized subspecies of genus Conopophaga (Whitney, 2003). Melanopareia maranonica was used as outgroup following Moyle et al. (2009).

DNA was extracted from tissues (muscle or blood) following Bruford et al. (1992). For most samples, we sequenced three mitochondrial genes – cytochrome b (cytB), NDH dehydrogenase subunit 2 (ND2) and NDH dehydrogenase subunit 3 (ND3); and two nuclear introns – glyceraldehyde-3-phosphate dehydrogenase intron 11 (G3PDH) and transforming growth factor beta 2 intron 5 (TGFβ2). Primers used to amplify these genes are given in Table S2 in supplementary material. PCR (25 μL) contained template DNA (50 ng), 1X of Taq buffer (GE Healthcare), dNTPs (0.32 μM), 0.5 μM of each primer and 0.5 U of Taq polymerase (GE Healthcare). PCR conditions were as follows: an initial denaturation step at 94 °C for 3 min and 30 s; followed by 35 or 40 cycles at 94 °C for 35 s, annealing temperature for 40 s and 72 °C for 1 min; plus a final extension step at 72 °C for 9 min. Annealing temperatures were as follows: cytB, ND3, and ND2-56°C; G3PDH-63°C; and TGFβ2-60°C.

Amplicons were purified using polyethylene glycol 20% (PEG) precipitation (Sambrook et al., 1989) or a mixture of Exonuclease 1/FastAP (Thermo Scientific). DNA was directly sequenced in both directions with the same amplification primers using Big Dye terminator 3.1 cycle sequencing kit (Applied Biosystems) following the manufacturer’s protocol. Sequences were analyzed in an automated sequencer ABI PRISM 3100 (Applied Biosystems). Some samples were sequenced at the Macrogen sequencing service (Macrogen Inc., Seoul, South Korea).

Electropherograms were inspected and assembled in contigs using CodonCode Aligner v. 3.7 (CodonCode Inc.). Heterozygous sites in nuclear introns were coded according to IUPAC code when double peaks were present in both strands of the same individual’s electropherograms. Sequences were aligned using the CLUSTAL W method (Higgins et al., 1994) in MEGAS (Tamura et al., 2011). All alignments were inspected and corrected visually.

#### 2.2. Phylogenetic analyses

We used Bayesian inference (BI) and maximum likelihood (ML) to estimate the phylogeny of Conopophagidae based on a concatenated sequence matrix of five partitions (ND3, ND2, cytB, G3PDH, and TGFβ2). The best fit model for each gene was selected using MrModeltest 2.2 (Nylander, 2004) based on the Akaikie information criterion (AIC), in conjunction with PAUP* (Swofford, 1998). Phylogenies were estimated using MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) and RAxML (Stamatakis et al., 2008) for BI and ML, respectively. Two independent Bayesian runs of 20 million generations with four chains of Markov chain Monte Carlo (MCMC) each were performed. The first million generations were discarded as burn-in, after which trees were sampled every 500 generations. Chain convergence (Effective Sample Size – ESS values > 200) was checked using the likelihood plots for each run using Tracer 1.5 (http://beast.bio.ed.ac.uk/Tracer). The Potential Scale Reduction Factor was also used to check chain convergence and burn-in; values close to one indicate good convergence between runs (Gelman and Rubin, 1992). RAxML analysis was run under the GTRAC model; invariable sites and gamma distribution were estimated for each partition during the run. Node supports of the maximum likelihood analyses were estimated by 1000 bootstrap replications. In addition, we obtained trees separately in RAxML for mitochondrial (ND3, ND2 and cytB) and nuclear (G3PDH and TGFβ2)
alignments. MrBayes and RAxML analyses were carried out at CIPRES Science Gateway (Miller et al., 2010).

2.3. Divergence time estimation

Divergence time estimates were obtained by implementing a Bayesian relaxed clock model in BEAST 1.7.4 (Drummond et al., 2012) based on a concatenated mitochondrial dataset as well as all genes concatenated, and using the CIPRES Science Gateway. We used relaxed clock with an uncorrelated lognormal distribution (Drummond et al., 2006) as indicated by hLRT tests, UPGMA starting tree and Yule process for both datasets. We considered the same partitions and models used in the phylogenetic reconstructions, as estimated in MrModeltest 2.2. As the fossil record of

Fig. 1. (a) Phylogenetic relationships of Conopophagidae. The topology was obtained by Bayesian inference based on 3309 bp of mitochondrial and nuclear gene concatenated sequences. Node supports are posterior probabilities (PP) and bootstrap (BT) values for Bayesian inference and maximum likelihood, respectively. Asterisks indicate when both PP and BT were maximum (1.0 and 100, respectively). The colors and codes in the tree refer to the maps. (b and c) Maps with geographic distribution of Conopophaga species. Circles on the map indicate sampling sites and their codes correspond to the sequences in the tree. Species distributions are based on digital maps provided by Ridgely et al. (2003). The gradient of gray color in the map represents the elevation gradient (the darker the higher is the altitude). Conopophaga illustrations are courtesy of Lynx Edicions (Handbook of the birds of the world, Vol. 8, 2003). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
New World suboscines is very sparse, we used two different strategies of calibration to obtain absolute divergence times: (i) concatenated dataset – a secondary calibration of the root of Conopophagidae family as estimated by Olsson et al. (2013) of 20.05 Mya (95% confidence interval: 15.66–24.86) under a normal distributed prior; (ii) mitochondrial dataset – a mutation rate of mitochondrial genes (under a normal distributed prior) for birds of 1.05% (±0.05) per lineage per million years (Weir and Schluter, 2008). We performed two independent runs with each dataset with 100 million generations each, with parameters sampled every 10,000 steps and a burn-in of 10%. We checked for convergence between runs and analysis performance using Tracer 1.5, and accepted the results if ESS values were >200. The resulting trees were combined in TreeAnnotator and the consensus species tree with the divergence times was visualized in FigTree 1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

2.4. Ancestral area reconstruction

We used the package BioGeoBEARS (BioGeography with Bayesian Evolutionary Analysis in R Scripts; Matzke, 2013; http://cran.r-project.org/web/packages/BioGeoBEARS/index.html) to infer the ancestral areas for the Conopophaga clade. This package uses a maximum likelihood method similar to LAGRANGE (Ree et al., 2005; Ree and Smith, 2008). In this inference, ancestral areas are optimized onto internal nodes. BioGeoBEARS calculates maximum likelihood estimates of the ancestral states (range inheritance scenarios) at speciation events by modeling transitions between discrete states (biogeographical ranges) along phylogenetic branches as a function of time. It allows the use of the LAGRANGE DEC model (Dispersal Extinction Cladogenesis) and a new model called BioGeoBEARS DEC + J model (see Ree et al., 2005 and Matzke, 2013 for further details). Both models include two free parameters (d = dispersal and e = extinction), but DEC + J includes the additional parameter j that corresponds to founder event speciation. Likelihood values of these models were compared using Likelihood Ratio Test (LRT). We defined four geographical areas for the BioGeoBEARS analysis after considering the evidence available for historical relationships between relevant geographic areas in South America (Hoorn et al., 2010) and the distribution of Conopophaga taxa. These regions were as follows: Andes, South Amazonia, North Amazonia, and Atlantic Forest. Two of the most stable geological regions in South America are represented in these areas: Brazilian Shield (South Amazonia and Atlantic Forest), and Guiana Shield (North Amazonia). The ancestral area probability was computed for each node and subsequently plotted on the majority-rule chronogram using R scripts (available from P.-H.F. on request). We set the maximum number of areas at four.

2.5. Diversification rate analysis

We used R (R Development Core Team, 2012) with the APE and the LASER packages (Paradis et al., 2004; Rabosky, 2006) to calculate the diversification rate of genus Conopophaga. The analyses were computed on the maximum clade credibility tree and on 1000 phylogenies randomly sampled from the posterior distributions of trees (excluding the burn-in) in order to take into account the phylogenetic uncertainty. We kept only one lineage per species as delineated by our molecular dating results. We tested for constant diversification rates over time using both the γ statistic (Pybus and Harvey, 2000) and the maximum likelihood method of the AICRC (Rabosky, 2006). We obtained the null distribution for both estimations from 5000 simulated topologies of the same sample size. The sample size used (number of species; N = 11) was based on the results from the molecular species phylogeny. To test the effects of undetected or extinct species, we used three further levels of total species richness, by assuming that the known number of species represents 75%, 50%, and 25% of the true number of species. The trees were simulated based on the assumed total species number, and tips were reduced to the number sampled in our phylogenies. All our simulations assumed that missing species were randomly distributed in the phylogeny. We assumed that our taxon sampling of Conopophaga is complete as all except two of the currently recognized subspecies were analyzed and most of these taxa were represented by samples from multiple sites. A strongly negative γ is associated with a decrease in diversification rates over time (rejection of constant diversification), whereas a positive value could represent constant or increasing diversification rates. We used a one-tailed test to detect significant negative γ values, which is considered conservative if extinction is nonzero (Pybus and Harvey, 2000). We subsequently fitted five maximum-likelihood diversification models; (i) the Yule model (constant speciation rate without extinction), (ii) a birth–death model (constant speciation rate and constant, nonzero extinction rate), (iii) a diversity-dependent diversification model with exponentially decreasing speciation, (iv) zero extinction rates, (v) a modified Yule model allowing for two different speciation rates with a breakpoint (Rabosky, 2006). The latter three represent models with rates that vary through time with a diversity-dependent diversification model with linearly decreasing speciation and zero extinction rates. The AICRC measure is computed as the difference in AIC values between the best rate-variable model and the best rate-constant model. The AICRC is positive if the best model is rate-variable. Significant positive AICRC values were implemented using a one-tailed test against the simulated null distributions (Rabosky, 2006).

Using GEIGER (Harmon et al., 2008) and apTreeshape packages (Bortolussi et al., 2006) we implemented two approaches to identify nodes that display significant shifts in diversification rates: the Δ1 statistic, which considers topological information only, and the relative cladogenesis test, which tests lineages within time slices along the whole chronogram for differences in the number of descendants (Nee et al., 1992).

3. Results

3.1. Phylogenetic inferences

We obtained a concatenated matrix of 3309 bp for most samples (46 individuals, including outgroup): 343 bp of ND3, 941 bp of ND2, 986 bp of cytb, 391 bp of G3PDH, and 648 bp of TGF2B. Considering the ingroup only, there were 129, 382, and 357 variable sites in ND3, ND2, and cytb, respectively. No indels, unexpected stop codons, or ambiguous peaks in the electropherograms were found in these sequences, suggesting that they were of mitochondrial origin. In the nuclear intron sequences, we observed 59 and 36 variable sites in G3PDH and TGF2B, respectively. For the alignment of G3PDH we found five indels that varied between one to eight bp long. Three of these are congruent with the phylogenetic results: an indel of eight bp was shared by samples of C. lineata vulgaris, an indel of four bp was found only in C. aurita australis, an indel of one bp was shared only by samples of C. melanogaster, and another indel of one bp was found only in C. ardesiaca. The remaining indel of one bp was found in C. I. ceareae, C. melanogaster, C. melanops and C. roberti, and its presence was not congruent with the phylogeny. We found three indels in TGF2B (two of six bp, and one of one bp); one indel of six bp was present in all samples of C. melanops and C. melanogaster, a six bp indel was shared between C. aurita snethlageae and C. aurita pallida, and a one bp indel was shared between C. a. australis and C. a. occidentalis. These indels in TGF2B were all congruent with the phylogenetic results (Fig. 1), thus providing additional support for these clades. All sequences were deposited in GenBank (accession numbers, ND3: KJ912712.
indicated a decrease of diversification rate through time

Furthermore, all species were monophyletic with the exception of *C. lineata*, as *C. lineata cearensis* did not cluster with *C. l. lineata* and *C. l. vulgaris* (Fig. 1). The BI tree recovered *C. l. cearensis* as sister of *C. peruviana* (Fig. 1), but with low node support (posterior probability [PP] of 0.93). Also, the relationships of *C. l. cearensis* with other species could not be resolved in the ML phylogeny. The ML tree for the nuclear dataset recovered most species as monophyletic, but the relationships between them were poorly supported (not shown).

Two major clades were observed in the phylogeny: one comprising genus *Pittasoma*, and another, *Conopophaga* (Fig. 1). Within *Conopophaga*, three well supported major clades were recognized. A basal lineage was comprised by the southern Amazonian *C. melanogaster*, while the other cluster included a clade with *C. melanops* and another clade with all remaining *Conopophaga* species (Fig. 1). Basal relationships within this latter clade were poorly resolved and are best regarded as forming a polytomy. Nevertheless, some nodes in this clade showed strong support (Fig. 1): (i) *C. ardesiaca* as sister of *C. castaneiceps*; (ii) *C. lineata* as sister of *C. roberti*.

Our phylogenies also revealed most subspecies as monophyletic, but within *C. ardesiaca* the subspecies *C. a. saturata* was paraphyletic with respect to *C. a. ardesiaca*. Within *C. aurita* and *C. lineata* we also found deep divergences between subspecies (Fig. 1). The form *C. l. cearensis* represents an independent lineage (see above). Our phylogenies also revealed genetic structure within *C. l. vulgaris*, with two clades that are congruent with a central (Fig. 1; L1 and L2) and a southern (Fig. 1; L3, L4 and L5) Atlantic Forest distribution. *C. aurita* exhibited a structured diversity in western Amazonia (*C. a. australis* and *C. a. occidentalis*) and eastern Amazonia (*C. a. snethlageae* and *C. a. pallida*). Also, in *C. melanops* three clades were recovered which correspond each to recognized subspecies (Fig. 1): *C. m. nigrifrons* (M7 and M8), *C. m. perspicillata* (M6), and *C. m. melanops* (M1, M2, M3, M4 and M5).

3.2. Divergence times

The absolute divergence times of the gnatelers generated by BEAST revealed an initial divergence of *Pittasoma* and *Conopophaga* during the early Miocene and the radiation within *Conopophaga* during the late Miocene (Fig. 2). Divergence time estimates for concatenated and mitochondrial datasets were very similar for all nodes within *Conopophaga*, and all 95% of high posterior densities (HPD) for node dates overlapped between these two estimations (Table 1). The oldest divergence was the split between *Pittasoma* and *Conopophaga* (~18 Mya; Fig. 2; Table 1). Within *Pittasoma*, the divergence between its two species was ~6 Mya (Fig. 2; Table 1). Within *Conopophaga*, the oldest divergence was the basal split between *C. melanogaster* and the remaining species (~10.5 Mya), whereas the most recent was the split between *C. ardesiaca* and *C. castaneiceps* (~3 Mya; Fig. 2; Table 1). In addition, comparatively old divergences (~6.5–1.5 Mya) were also found between subspecies in some of the polytypic species (i.e. *C. aurita*, *C. lineata*, and *C. castaneiceps*).

3.3. Ancestral distribution of *Conopophaga*

The model that best fitted the phylogenetic reconstruction of *Conopophaga* was DEC + J (ln L = −40.3), while DEC (ln L = −44.1) was less likely (LRT = 7.2, p < 0.05). The most likely ancestral area reconstruction is given in Fig. 2 (see Fig. S1 in supplementary material for node pie charts likelihoods of ancestral areas). This biogeographic reconstruction indicated a South Amazon plus Atlantic Forest origin, which comprises one of the geologically most stable areas in South America, the Brazilian Shield. Our analysis suggested a vicariant event during the Late Miocene, with *C. melanogaster* and *C. melanops* in opposite sides of the shield, in the southern Amazon and the Atlantic Forests. The most likely origin for the clade containing the remaining *Conopophaga* species (node D), except *C. melanogaster*, was the Atlantic Forest. From this area, the clade that holds the remaining *Conopophaga* species seems to have recolonized the Amazon area. Other putative major events involved: (i) the dispersal of the ancestor of *C. castaneiceps* and *C. ardesiaca* (node I) toward the Andes and their speciation possibly due to the orogenic changes in the eastern Andean region; (ii) two vicariant events between the Atlantic Forest and the Amazon involving the splits between *C. lineata* and *C. roberti* (node H), and *C. l. cearensis* and *C. peruviana* (node F); (iii) the range expansion of *C. aurita* to the North Amazon (node G). However, it is noteworthy to remind that this ancestral area reconstruction should be interpreted with caution, as some nodes were poorly resolved in our phylogenetic reconstructions (Fig. 1).

3.4. Diversification rate

The lineage-through-time plot for *Conopophaga* (Fig. 3) levelled off slightly after an initial rapid diversification, generating an apparent diversity-dependent pattern. Both the γ statistic and AICRC indicated a decrease of diversification rate through time and the diversity dependent model being the best model for subspecies diversification (Tables 2 and 3). The observed γ and AICRC of the maximum clade credibility tree for *Conopophaga* was significantly lower than expected if diversification were constant through time even if the tree only included 50% of the true diversity. We did not identify any shifts in diversification rates in *Conopophaga* (Delta 1 or RRT test).

4. Discussion

4.1. Origin and diversification of the gnatelers in South America

Although *Conopophagidae* represents one of the oldest lineages among the New World suboscines (~41 Mya; Ohlson et al., 2013), this family has very few species when compared to other lineages of similar age within the New World suboscines, such as the Thamnophiliae, Furnarioidae and Tyrannoidae groups. This suggests that extinction of deep branches may have occurred, as was also observed by the imbalance of branching in many parts of the Suboscines chronogram by Ohlson et al. (2013). Besides, it is worth mentioning that the radiation of the *Conopophaga* clade started rather recently (upper Miocene; ~11 Mya; Fig. 2). Thus, *Pittasoma* seems to be a possible relicual lineage that diverged during the late Oligocene to early Miocene (Fig. 2; Table 1). Moreover, the imbalanced phylogeny of the Thamnophilidae superfamily, which contains old and species-poor genera, such as *Melanopareia* (*Melanopareiaeidae*), *Pittasoma*, and *Euchrepomis* (*Euchrepominae*), seems to suggest a relicual pattern, where extinctions may have put a veil over the early radiation of thraupine suboscines.

Despite this, the predominance of taxa representing deep Thamnophilidae lineages in the Eastern Brazilian-Andean area (Irestedt et al., 2002) may suggest an origin in this region. *Conopophaga* species are mainly distributed within the Brazilian Shield area and along the western margin of the Amazon Basin, corresponding to an Eastern Brazilian-Andean biogeographical pattern
It is in accordance with our ancestral area reconstruction, which indicated that the ancestor of the genus was probably distributed around the Brazilian Shield (Silva, 1995). Thus, only the widespread Amazonian C. aurita extends north of the Amazon, in Amapá and the Guianas, and up to Rio Negro, which could represent an expansion before the establishment of the Amazonas river drainage channel to the Atlantic Ocean in the Plio-Pleistocene (Campbell et al., 2006; Latrubesse et al., 2010). Notwithstanding, as our analyses do not include samples of C. aurita from north of the Amazon River, our statement of a Brazilian Shield origin for Conopophaga could change if the population from that region would not cluster with any of the lineages of C. aurita recovered in our study.

The dating of the split between the Eastern Brazilian-Andean Conopophaga and Pittasoma of the Chocó region (Fig. 2; node A, ~18.8 Mya) agrees with the early mountain building in the northern Andes (Hoorn et al., 2010: 23–10 Mya), when the biota of the northern proto-Andean peninsula (corresponding to the Chocó region) was isolated from that of Amazonia. The young divergence between P. rufopileatum and P. michleri can be attributed to a
recent dispersion after the closure of the Panama isthmus by an ancestor of *P. michleri*. This is supported by the confidence intervals of our time estimates (Fig. 2; Table 1) and is in accordance with the estimated geological period of the Panama isthmus closure (~4–3 Mya; Coates and Obando 1996; Kirby et al., 2008). Other passerines exhibited a similar pattern of dispersion (Smith and Klicka, 2010; Weir et al., 2009).

The ancestral area reconstruction of *Conopophaga* could indicate an origin along the margins of the Brazilian Shield (Figs. 2 and 4). In our biogeographic inference the ancestral lineage of *Conopophaga* seems to have its distribution centered on the Brazilian Shield and started to colonize the western Amazonia basin only as the Andes uplift (Hoorn et al., 2010), marine incursions (Antonelli and Sanmartín, 2011; d’Horta et al., 2013; Patel et al., 2011). Our diversification result thus corroborates the relictual status of *Conopophaga*, which seems to have filled a limited ecological space in the South America (a diversity-dependent pattern). This pattern might also be linked to the very specialize ecology of *Conopophaga*, which have in consequence restricted speciation opportunity compared to other New World suboscines. Moreover, we did not detect any significant shift in the diversification rate (see results section), a fact against a potential growing scenario. In such a growing scenario, the origin of Neotropical biota cannot be attributed to only one or few events during key time intervals (Rull, 2011a). This result reinforced the idea that *Conopophaga* represent an old and very specialize lineage, less prone to important diversification compared to its New World suboscines relatives. Despite this poor diversity, our Neotropical *Conopophaga* phylogeny provides a good biogeographical model shaped by a complex sequence of landscape changes through the Miocene and the Quaternary (see previous discussion part, Fig. 4). These can include Andes uplift (Hoorn et al., 2010), marine incursions (Antonelli and Sammartin, 2011; Hoorn et al., 2010), Plio–Pleistocene formation of flat depositional basins with the

Table 1
Divergence times and confidence intervals (95% of high posterior density (HPD)) in millions of years (Mya) between major clades of *Conopophagidae*. Node letters follow BEAST phylogeny. Model names as in laser (Rabosky, 2006). AIC scores are provided for each of the empirical LTTs. AIC scores from the rate-constant and rate-variable models providing the best fit are noted with bold, as well as AIC values and $P$-values.

<table>
<thead>
<tr>
<th>Node</th>
<th>Age Mya (95% of HPD)</th>
<th>Concatenated</th>
<th>Mitochondrial</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>9.27 (12.53–6.55)</td>
<td>9.73 (11.38–8.08)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>6.97 (9.37–4.89)</td>
<td>7.80 (8.45–6.26)</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>6.62 (9.30–4.38)</td>
<td>5.43 (7.11–3.90)</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>6.18 (8.52–4.34)</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>5.09 (6.93–3.43)</td>
<td>5.24 (6.24–4.32)</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>3.03 (4.24–2.01)</td>
<td>3.14 (3.85–2.45)</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>3.04 (4.31–1.94)</td>
<td>3.27 (4.06–2.38)</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>2.74 (3.83–1.83)</td>
<td>2.85 (3.49–2.25)</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>2.65 (3.75–1.71)</td>
<td>2.64 (3.30–1.99)</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.33 (3.29–1.50)</td>
<td>2.44 (3.07–1.83)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>2.04 (2.91–1.29)</td>
<td>2.10 (2.65–1.54)</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>1.29 (1.90–0.78)</td>
<td>1.24 (1.65–0.87)</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>1.27 (1.89–0.78)</td>
<td>1.24 (1.62–0.90)</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0.89 (1.32–0.54)</td>
<td>0.9 (1.20–0.60)</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.89 (1.38–0.51)</td>
<td>0.88 (1.24–0.56)</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.67 (1.06–0.37)</td>
<td>0.7 (0.89–0.42)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Test of constancy of diversification rates using the $\gamma$ and AICrc statistics. Observed statistics are for the maximum clade credibility (MCC) trees; $P$-values (significant at $\leq 0.05$ in bold) are shown for the MCC tree and the 95 percentile of a random sample (1000 trees) from the posterior distribution of trees. The $P$-values were generated from simulated null distributions of 5000 trees for each taxon and each of the assumed total species numbers. Simulations all accounted for the number of species not sampled in our phylogenies, and assumed that all species were known (100%), or that only 75%, 50%, and 25% were known, respectively.

<table>
<thead>
<tr>
<th>$P$-values</th>
<th>Observed statistic</th>
<th>100% known</th>
<th>75% known</th>
<th>50% known</th>
<th>25% known</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCC 95th</td>
<td>MCC 95th</td>
<td>MCC 95th</td>
<td>MCC 95th</td>
<td>MCC 95th</td>
</tr>
<tr>
<td>GammaMCC</td>
<td>–1.34</td>
<td>0.04</td>
<td>0.05</td>
<td>0.09</td>
<td>0.16</td>
</tr>
</tbody>
</table>
| dAICrc     | 1.23               | 1.27       | 0.18      | 0.27      | 0.09      | 0.16      | 0.27      | 0.4       | 0.46      | 0.6
Amazonian megafans and establishment of the east-draining Amazon river (Ribas et al., 2012), as well as Pleistocene climatic-vegetational changes (Carnaval et al., 2009; Cheng et al., 2013).

4.2. Diversification of the Atlantic Forest lineages

Three independent evolutionary lineages of Conopophaga are endemic to the Atlantic Forest: C. melanops, C. lineata, and C. l. ceareae (Fig. 1). C. melanops is the oldest lineage, and diverged from all other species of Conopophaga in the late Miocene (Fig. 2; Table 1). Recent studies showed that the Atlantic Forest holds ancient lineages that date from the mid-Tertiary for birds (Derryberry et al., 2011), mammals (Fabre et al., 2013; Galewski et al., 2005) and frogs (Fouquet et al., 2012). C. lineata and C. l. ceareae originated around the late Miocene (Fig. 2; Table 1) and have Amazonian sister species, in accordance with hypotheses of historical connections between Amazonia and Atlantic Forest biota (Batalha-Filho et al., 2013; Cheng et al., 2013; Costa, 2003; Willis, 1992).

Our data confirmed the Atlantic Forest diversity, and revealed several cryptic intra-specific lineages. It is noteworthy to highlight distinct diversification patterns of C. melanops, C. lineata and C. l. ceareae. Indeed, C. lineata exhibited higher mean intra-specific uncorrected genetic distance based on mtDNA data (C. lineata, 3.51%; C. melanops, 0.72%; C. l. ceareae, 0.58%), as well as longer intra-specific branch lengths (Figs. 1 and 2). In C. lineata the first divergence occurred during the late Pliocene to early Pleistocene, whereas in C. melanops and C. l. ceareae the split took place in the Pleistocene. Interestingly, C. lineata and C. melanops have similar latitudinal distribution in the Atlantic Forest (Ridgely and Tudor, 1994; Fig. 1) and similar geographic structure (two components: one in the north and one in the south), but they exhibited distinct temporal splitting patterns (Fig. 2; Table 1). This incongruence on the temporal patterns could be related to the different elevation distributions: C. lineata inhabits upland forest (500–2400 m) while C. melanops is mainly found in lowland forest (up to 800 m) (Whitney, 2003). Lowland and upland regions in the Atlantic Forest...
seem to have been affected by the Quaternary climate changes in different ways. In general, species of lowland forests show strong signatures of the effect of the last glacial maximum (Carnaual et al., 2009), whereas some upland organisms do not (Amaro et al., 2012; Batalha-Filho et al., 2012).

4.3. Diversification in the Amazon basin

While the first split of Amazonian gnateaters, which gave rise to *C. melanogaster*, is of mid-Miocene age, the other three splits occurred through the late Miocene to early Pliocene (Fire 2; Table 1). As in the Atlantic Forest, the four Amazonian species (*C. melanogaster*, *C. aurita*, *C. peruviana*, and *C. roberti*) represent independent evolutionary lineages in the phylogeny (Fig. 1). However, it is important to note that, because of the poor resolution of some deep branches, some Amazonian species, such as *C. aurita* and *C. peruviana*, could in fact present a sister taxa relationship. The lack of monophyly of the Amazonian species was also observed in other forest birds, such as *Brotogeris* (Ribas et al., 2009) and *Sclerurus* (d’Horta et al., 2013). Thus, it is possible that the Amazonian biota may have multiple origins (Ribas et al., 2009) and *Conopophaga* may represent a new example of expansion from the East Brazilian-Andean elevated land surfaces.

Intra-specific diversification within Amazonia was detected only in the geographically widespread *C. aurita* (Fig. 1), which also shows the largest number of recognized subspecies (Whitney, 2003). The most basal divergence within *C. aurita* dates back to late Miocene to early Pliocene (Fire 2; Table 1), and divides western and eastern Amazonian subspecies. Within these two groups further diversification took place in the Pliocene (Figs. 1 and 2). Given the possible origin of *Conopophaga* in the Brazilian Shield, we might assume a past expansion of *C. aurita* from this area before the final establishment of the Amazonian drainage channel to the Atlantic Ocean (through Late Miocene and Pliocene), and when the Japurá-Negro palaeochannel formed the northern fan margin of the central Amazon basin (Wilkinson et al., 2010). Further studies with a denser taxon sampling, including samples from the north of the Amazon River are needed to understand the phylogeographical structure of *C. aurita*.

4.4. Uplift of the Andes and diversification of gnateaters

The Andean species, *C. castaneiceps* and *C. ardesiaca*, form a monophyletic group with corresponding northern and central clades, respectively (Fig. 1). The Andean clade originated around the Late Miocene (Fire 2), and the split might be related to the transitions between the lowland Amazon and the Andes (Upham et al., 2013). Divergence between Andean species occurred in the Pliocene (Fire 2; Table 1). They overlap marginally in Cuzco, eastern Peru (Schulenberg et al., 2010), in a region with many species replacements, in the transition between the perhumid tropical Andes and the more seasonal “yungas” forest of eastern Peru and Bolivia (Bonaccorso, 2009; Gutiérrez-Pinto et al., 2012; Isler et al., 2012). We also detected divergent lineages in *C. castaneiceps* that split in the late Pliocene (Fig. 1 and 2). This pattern of diversification in the Andes is possibly correlated to the mountain building pattern of the Cordillera, that follows a south to north orogeny (Hoorn et al., 2010), and is in accordance with the rate of cladogenesis observed in other avian groups (Chaves et al., 2011; Ribas et al., 2007).

4.5. Systematics

In this study we provided the first comprehensive molecular phylogeny of the Neotropical gnateaters. Our phylogeny strongly supported both *Conopophaga* and *Pittaomus* as reciprocally mono-

phyletic lineages (Fig. 1), and also provided evidence for the taxonomic revision of *Conopophaga*. Even though relationships within *Conopophaga* were not fully resolved (Fig. 1), nine clades were well supported, including all previously recognized species and a paraphyletic relationship in *C. lineata*. The phylogeny placed the Andean taxa *C. castaneiceps* and *C. ardesiaca* as sister species, while the Atlantic Forest species were not monophyletic (Fig. 1). The form *cearae* has, until now, been recognized as a subspecies of *C. lineata*, based on plumage phenotypes (Whitney, 2003), but our results strongly suggests *C. l. cearae* as a separate species, and our Bayesian concatenated phylogeny recovered it as sister of *C. peruviana*, although with low support (0.93 of PP). Better genetic and geographical sampling of the disjunct geographical populations of *C. l. cearae* (in northern Chapada Diamantina, patches of forest amidst dry Caatinga locally known as Brejos de Altitude, and northern São Francisco River) may help to reveal ancient connections between Amazonia and the Atlantic Forest. The similarities in plumage between *C. lineata* and *C. (lineata) cearae* could be the result of retention of ancient polymorphism or parallel evolution, as these species share very similar plumage phenotypes.

We also observed genetic structure in four species (Fig. 1): *C. melanops*, *C. castaneiceps*, *C. lineata*, and *C. aurita*. As some of these clades are phenotypically and vocally diagnosable (Whitney, 2003), a detailed taxonomic revision of the genus may support that several of these subspecies also could be ranked as full species. In addition, further phylogeographic studies of these species with dense population sampling will help to comprehend the geographic variation in these taxa.

4.6. Conclusions

Our analyses of the diversification of the Conopophagidae family provided an additional scenario for the evolution in *Terra Firme* forest across tropical South America. The spatio-temporal pattern may reflect an evolutionary history influenced by several independent external events such as the Andean uplift (Hoorn et al., 2010), marine incursions (Antonelli and Sanmartin, 2011), paleoclimatic events (Carnaual et al., 2009; Cheng et al., 2013) as well as river formation (Ribas et al., 2012). Nevertheless, it is intriguing to observe that the *Conopophaga* clade diversified so recently. Apparently this diversification took place only within elevated/erosional landscapes, primarily south of the Amazon river, and outside the flat sedimentary parts of the wetlands and megafans of the western and central Amazon basin. Other old South American radiations took place within the Brazilian and Guiana shields, with recent Plio-Pleistocene radiation driven by wetland dynamics within the basin. However, *Conopophaga* may have been geographically more limited and did not show the same strong radiation (driven by river barriers) within the floodplains (with some exception for *C. aurita*).

Therefore, the origin of the Neotropical region biota seems to be a mosaic of many complex evolutionary scenarios (Rull, 2013), and as more studies become available more detailed scenarios may arise, and a more complete history will be elucidated in the future.

Acknowledgments

We thank John Bates (FMNH) and Nate Rice (ANSP) for providing some of the tissues used in this study. We thank Fernando M. d’Horta, Renato G. Lima, Gustavo S. Cabanne, and Guilherme R. Brito for collecting some samples used in this study. Amy Chernaksky from Lynx Edicions kindly provided permission to use images from Handbook of Birds of the World. We thank Gustavo Bravo for suggestions on previous version of the manuscript. We thank an anonymous reviewer and the Editor Carey Krajewski for
their comments. This study was co-funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (2009/12989-1, BIOTA 2013/50297-0), NSF (DOB 1343578), NASA, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). JF and PHF thanks the Danish National Research Foundation for funding the Center for Macroecology, Evolution and Climate; PGPE and MI thanks the Swedish Research Council for funds (Grant No. 621-2010-5321 to P.G.P.E.). PHF was supported by Marie-Curie grants (PIOF-GA-2012-330582-CANARIP-RAT, FP7 CIG-293845). Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) and Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) provided permits to collect the samples. This work was developed in the Research Center on Biodiversity and Computing (BioComp) of the Universidade de São Paulo (USP), supported by the USP Provost’s Office for Research.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2014.06.025.

References


