Diversity, biogeography, and reproductive evolution in the genus Pipa (Amphibia: Anura: Pipidae)

Antoine Fouquet, Josselin Cornuault, Miguel T. Rodrigues, Fernanda P. Werneck, Tomas Hrbek, Andrés R. Acosta-Galvis, David Massemin, Philippe J.R. Kok, Raffael Ernst

Abstract. The genus *Pipa* is a species-poor clade of Neotropical frogs and one of the most bizarre-looking due to many highly derived anatomical traits related to their fully aquatic life-style. With their related African lineages, they form the Pipidae family, which has attracted much attention, especially regarding its anatomy, reproductive biology, paleontology and biogeography. However, the actual diversity and relationships within *Pipa* remain poorly understood, and thus, so do their historical biogeography and the evolution of striking features, such as the absence of teeth and endotrophy in some species. Moreover, phylogenomic results display substantial temporal incongruences suggesting substitution rate heterogeneity across pipids. Using short mtDNA sequences across the distribution of the genus we identified 15 OTUs, a ~~two~~ twofold increase compared to the seven nominal species currently considered valid. Some of the closely related OTUs do not share nuDNA alleles, confirming species divergence. Time calibrated phylogenies obtained from mitogenomes and from 10 nuclear loci provide highly similar topologies but strikingly distinct node ages for *Pipa*. We identified high dN/dS ratios along the mtDNA stem branch of *Pipa*, probably because an acceleration of the substitution rate in *Pipa* increases the effect of saturation on fast evolving positions. Focusing on the nuDNA topology, we hypothesize an early Neogene amazonian origin of the diversification of *Pipa*, with an initial split between the Guiana-Brazilian Shields and Western Amazonia, a pattern found in many other codistributed groups. All the western species are edentate, suggesting a single loss in the genus. Each of these groups diversified further out of Amazonia, toward the Atlantic forest and toward Trans-Andean forest, respectively, and these events are concomitant with palaeogeographic changes and recovered in codistributed groups. The two amazonian lineages have probably independently acquired endotrophic larval development.

Keywords. Amazonia; Atlantic Forest; Endotrophy; Neogene; Species delimitation

1. Introduction

The increasing accumulation of genomic data has permitted unveiling phylogenetic relationships and divergence times with unprecedented accuracy throughout the tree of life (Delsuc et al., 2005; Burki et al., 2020). This is the case for amphibians in which phylogenomic investigations have spectacularly improved our understanding of the relationships among major lineages (e.g. Neobatrachia, Streicher et al., 2018), and, combined with fossil calibrations, dramatically narrowed down temporal estimates of their origins (Feng et al., 2017; Hime et al., 2020). These progresses also notably led to the perception that mtDNA, which has been used for decades to decipher evolutionary patterns, sometimes underwent rapid and important shifts in substitution rates (Irisarri et al., 2012; 2017), sometimes leading to long-branch attraction artefacts (Gissi et al., 2006; Irisarri et al., 2017). The impact of the variations in substitution rate across loci and along the branches on phylogenetic reconstructions has been recognized for a long time and methods such as relaxed clocks implemented in BEAST (Drummond and Rambaut, 2007; Drummond et al., 2012; Bouckaert et al., 2014) are specifically designed to take into account these variations in order to provide more reliable time estimates. However, the reliability of these methods still largely depends on the availability/quality of priors such as node age calibration, for example from fossils, that are not evenly distributed along the tree of life and are simply absent in many groups (Donoghue and Yang, 2016).

The frogs of the family Pipidae are the only ones with a fully aquatic lifestyle, and harbor many highly derived anatomical traits (Cannatella, 2015) and chromosomal features (Mezzasalma et al., 2015). Extant taxa of the family Pipidae include *Pipa* (7 species) in the Neotropics, and *Xenopus* (29 species, *Silurana* being considered a synonym), *Hymenochirus* (4 species), and *Pseudhymenochirus* (1 species) in Africa. The sister-taxon of Pipidae is Rhinophrynidae, with a single extant fossorial species, *Rhinophrynus dorsalis*, which occurs in central and southern North America. They form altogether the Pipoidea clade, an ancient group of frogs whose crown age is c.a. 160 Ma old (e.g., Hime et al., 2020). One pipid in particular, *Xenopus laevis*, has become a model organism and the focus of a tremendous amount of medical and fundamental lab research (e.g., Vleminckx, 2018). The relationships among pipids have also been the focus of many studies (e.g., Irisarri et al., 2011; Bewick et al., 2012; Hedke et al., 2013; Cannatella et al., 2015) including node and tip dating based upon what is probably the most abundant fossil record of all amphibian families (e.g., Baez and Pugener, 2003; Trueb et al., 2005; Gomez, 2016). At the family level, this group has attracted much attention not only because of their bizarre anatomical features, but also because of their biogeographic history tightly linked to the break-up of western Gondwana. Phylogenomic analyses (Hedke et al., 2013; Feng et al., 2017; Irisarri et al., 2017; Hime et al., 2020), either alone or in combination with morphological data and fossils (Cannatella, 2015), largely support the monophyly of African pipids which form the sister group of *Pipa*. Time estimates for the divergence between *Pipa* and the African pipids obtained from phylogenomic studies are highly congruent, and generally predate the final stage of the break-up of western Gondwana (c.a. 105 Ma) (Feng et al., 2017; Irisarri et al., 2017; Hime et al., 2020). Nevertheless, an alternative topology, i.e., *Pipa* + Hymenochirini, has found support in a few genomic analyses (Bewick et al., 2012; Cannatella, 2015) and mitogenomic data alone (Irisarri et al., 2017; Zhang et al., 2021). Even though these studies included less genomic data than the most recent studies recovering *Pipa* vs. African pipids, this alternative topology is in line with the morphological similarity between *Pipa* and Hymenochirini (Cannatella and Trueb, 1988; Gomez and Perez-Ben, 2019). Given (1) the bias inherent to mtDNA and (2) the consensus provided by the most recent and most extensive phylogenomic investigations, the hypothesis of the monophyly of african pipids seems more robust, and thus either implies that the morphology of *Xenopus* is highly derived, and/or that *Pipa* and Hymenochirini underwent ~~convergent~~ evolution (Irisarri et al., 2011; Cannatella, 2015).

Paradoxically, despite the attention that pipids have attracted, the phylogenetic relationships within *Pipa* have been solely investigated using morphological characters (Trueb and Cannatella, 1986; Cannatella and Trueb, 1988; Trueb and Massemin, 2001), or based on short mtDNA sequences with incomplete taxonomic sampling (Vacher et al., 2020; Lima et al., 2020), and consequently remain poorly understood. Multilocus phylogenetic studies focusing on pipids have either included a single *Pipa* terminal or two, and provided highly inconsistent crown ages for the genus. Among these studies the ones that included the analysis of mitogenomes (Irissari et al., 2012; Evans et al., 2019; Hemmi et al., 2020) ~~found~~ crown ages > 50 Ma, which suggests ancient diversification. Conversely, Feng et al. (2017) using 88 kb of nuDNA loci found a Most Recent Common Ancestor (MRCA) for *Pipa* spp. to be only 11 Ma, implying a relatively recent diversification instead. As a corollary, Irissari et al. (2012) recovered disproportionate branch lengths between mtDNA and nuDNA within *Pipa* as compared to other non-Neobatrachia anurans. They also identified such imbalance in Neobatrachia and demonstrated that this was partly due to an acceleration of the substitution rate of the mtDNA, possibly resulting from a relaxed purifying selection on mtDNA. Therefore, incongruences among time estimates within *Pipa* could be due to a similar variation of mtDNA substitution rates. However, another paradox is the almost complete absence of known fossils directly related to *Pipa* (i.e., branching along the stem or nested within *Pipa*), which contrasts with an otherwise very rich Pipidae fossil record. The only exception is the c.a. 10 My old fossil from Corralito (lower Urumaco formation) of a portion of the sacrum of a large *Pipa*, with posteromedial ridges on the dorsal surface of each sacral diapophyses, suggesting close relationships with *Pipa pipa* (Delfino and Sanchez, 2018).

However, given its poor state of preservation, the phylogenetic position of this fossil remains unclear.

These knowledge gaps hamper an understanding of the biogeographic history as well as the evolution of reproductive modes and of the morphology of the genus. With four species occurring in Amazonia (*Pipa pipa* and *P. snethlageae* extend further), one in the Atlantic Forest (*P. carvalhoi*) and two with a Trans-Andean distribution (*P. parva* and *P. myersi*), one can reasonably assume that the genus originated in Amazonia. However, it remains conjectural to formulate any biogeographic hypothesis for *Pipa* in the absence of a robust time calibrated phylogeny, and since both ancient, i.e., 40–15 Ma (Fouquet et al., 2012a,b; 2013, 2014; Réjaud et al., 2020) and recent, i.e., < 5 Ma (Fouquet et al., 2014) dispersals events have been documented in anurans and in other vertebrates in the region (e.g. Ledo and Colli, 2017; Dal Vechio et al., 2018; Prates et al., 2018). It is also noteworthy that some characters are strikingly variable among *Pipa* spp., notably habitat and body size. *Pipa pipa* and *P. snethlageae* are large-bodied species (qualified as “macropipa”) and are distributed throughout Amazonia and even further into the Orinoco (Acosta-Galvis et al., 2016) and the Cerrado (Vaz Silva and Andrade, 2009; Dantas et al., 2019), occupying many different types of lotic and lentic aquatic environment, such as seasonally flooded forests. The other species like the Amazonian *P. aspera* and *P. arrabali* are small-bodied (“micropipa”), occupy various types of small water bodies, are occasionally terrestrial being able to colonize temporary ponds not connected to any rivers. *Pipa parva* also occupies temporary water bodies in the coastal deserts and semi-deserts of the Maracaibo basin and lowlands of the northeastern Caribbean region of Colombia (Galvis et al., 2011; Blanco et al., 2013). Considering these traits, we expect dispersal ability of the macropipa (e.g., da Fonte et al. 2021 reported *P. pipa* in floating meadows of the Amazon main course) to be high and their genetic structures to be relatively homogeneous over long distances vs. weaker dispersal abilities and more profound genetic structures in micropipa. Moreover, teeth are ~~only~~ present in *P. carvalhoi*, *P. aspera* (Trueb and Massemin, 2001) and in *P. arrabali* (Trueb and Cannatella, 1986), and although all the species of *Pipa* incubate their eggs in a specialized skin layer growing on the female's back after complex breeding behaviours (Rabb and Rabb, 1960; Weygoldt, 1976), only the amazonian species have a completely endotrophic larval development while the tadpoles of *P. carvalhoi*, *P. parva* and *P. myersi* hatch from the skin of the female and develop freely in the water instead (Greven, 2011). The evolution of edentulism (Paluh et al., 2021) and of reproductive mode (Vagi et al., 2019, Furness and Capellini, 2019) has recently attracted new attention thanks to the availability of extensive phylogenetic sampling and the development of new methods (Revell, 2012). However, like their biogeographical history, the monophyly of the edentate and of the endotrophic species has never been tested using molecular data, and the evolution of these characters remains ambiguous since one or the other of these characteristics necessarily impl~~y convergent~~ acquisition or loss.

In order to (1) delimit candidate species (2) evaluate phylogenetic relationships and divergence times within *Pipa* and (3) investigate biogeographic history, the evolution of reproductive mode and teeth loss/gain, we gathered mitogenomic and 10 nuDNA loci for all *Pipa* species and major Pipoidea lineages. Since temporal discrepancies have been previously recovered among pipids, we furthermore compared the substitution rates across mt and nuDNA and tested for variation in the dN/dS ratio across the genes and branches.

2. Materials and Methods

2.1. Species delimitation

Our first objective was to delimit major mtDNA lineages within *Pipa* since Vacher et al. (2020), Motta et al. (2018) and Lima et al. (2020) suggested that unrecognized species exist within *P. arrabali*, *P. aspera*, *P. pipa* and *P. carvalhoi*. Our sampling included 16S rDNA sequences from 115 specimens of *Pipa* (Supplementary table S1) covering the entire distribution of the genus. These samples were obtained through fieldwork and loans, and completed with available sequences from GenBank (Supplementary table S1). Newly acquired sequences were obtained from Sanger sequencing (details of primers are available in Supplementary table S5). DNA sequences were aligned on the MAFFT online server under the E-INS-i option with default parameters (Katoh et al., 2017) leading to a matrix of 595 base pairs (bp).

We applied three DNA-based single-locus species delimitation approaches using this matrix: (a) a distance-based method, the Automated Barcode Gap Discovery (ABGD; Puillandre et al., 2012), (b) a multi-rate coalescent-based method, the multi-rate Poisson Tree Processes model approach (mPTP; Kapli et al., 2017) and (c) a single-threshold coalescent-based method, the Generalized Mixed Yule Coalescent approach (single threshold GMYC; Pons et al., 2006; Monaghan et al., 2009). The ABGD delimitation was performed using the online web server (available at https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) with a prior of intraspecific divergences (K80) between 0.001 and 0.1 (P = 0.001–0.1), a proxy for minimum relative gap width of 1 (X = 1), and a number of steps equals to 30 (n = 30). For the mPTP delimitation, we first reconstructed a ML tree with RAxML v.8.2.4 (Stamatakis, 2014) using a CAT+Γ model which was estimated to be a suitable model via PartitionFinder V2.1.1 (Lanfear et al., 2017). The mPTP delimitation was undertaken on the tree rooted on *Rhinophrynus*, with 5 million Markov chain Monte Carlo (MCMC) iterations, sampling every 10,000th iteration and discarding initial 10% iterations as burn-in. For the GMYC delimitation, we reconstructed a time-calibrated phylogeny using BEAST 2.5.2 (Bouckaert et al., 2014). We used a birth and death population model to account for extinction processes and incomplete sampling. We used a single partition with a GTR+I+Γ substitution model, with an uncorrelated relaxed lognormal clock model of rate variation among branches (Drummond et al., 2006). We used the estimated age of the MRCA of Pipidae of Hime et al. (2020) as a calibration point, assuming normal prior distributions of 116.0 Ma (SD = 9 Ma). For the MCMC parameters, we used four independent chains of 100 million iterations, recording every 10,000th iteration. We combined the log and tree files of the four independent runs, discarding the first 30% iterations as burn-in, using LogCombiner 2.5 (Bouckaert et al., 2014) and checked the convergence of our parameters, confirmed by all ESS being above 200. We then extracted the maximum clade credibility tree (from 28,004 trees) using Tree annotator 2.5 (Bouckaert et al., 2014). We performed a GMYC delimitation on the ultrametric tree using the GMYC function of the {splits} R package (Ezard et al., 2009), with a threshold interval between 0 and 10 Ma and by using the single threshold method. Operational Taxonomic Units (OTUs) were defined using a majority-rule consensus from the results of the three methods, i.e., a lineage is considered as being an OTU if supported by at least two of the three methods.

2.2. nuDNA differentiation

We gathered sequences of four nuDNA loci (RAG1, POMC, BDNF, NCX1) for a subset of individuals in order to evaluate the degree of congruence of differentiation with the mtDNA-based delimitation (Supplementary table S2). Two OTUs (*P*. sp. “Negro”, *P*. sp. “Nordeste”) were not represented by any nuDNA sequences, one (*P. parva*) by only two nuDNA loci. Six OTUs were represented in all four datasets, but by a single terminal. Conversely, sequences from several individuals were obtained for six OTUs (only five OTUs for POMC). The sequences of NCX1 were incomplete in 3’ for five specimens or in 5’ for five others; therefore, we considered both subsets independently (NCX1a & b). We reconstructed Median-Joining networks (Bandelt et al., 1999) from PopArt 1.7 (Leigh and Bryant, 2015).

We considered the absence of nuDNA allele sharing among specimens assigned to closely related OTUs as indicative of congruent differentiation.

2.3. Molecular dating

We selected one representative of each delimited OTU, with the exception of two OTUs that were ~~only~~ represented by a limited amount of mtDNA and were discarded, for estimation of phylogenetic relationships and divergence times. We obtained whole mitogenomic data for representatives of 11 OTUs via shotgun sequencing (Supplementary table S3, GenBank accession numbers will be added upon manuscript acceptance; methodological details are available in Supplementary material S4). We completed the mtDNA matrix for the remaining two *Pipa* terminals for 12S, 16S, COI and Cytb using data available in GenBank (Table S4). We further complemented this mitogenomic dataset with 10 nuDNA loci (Supplementary table S4) via Sanger sequencing and sequences available in GenBank (details of primers are available in Supplementary table S5). During the matrix building, we noticed two sequences in GenBank from Irisarri et al. (2011) that appear to be swapped (AY341762=*Pipa parva*, AY341763=*Hymenochirus*), which were thus relabeled. We also retrieved from GenBank mitogenomes and homologous nuDNA sequences for five outgroups representing each African Pipidae genus and *Rhinophrynus dorsalis*, the sister group of Pipidae, including annotations that were transferred to the new mitogenomes. DNA sequences were realigned on the MAFFT online server under the E-INS-i option for 12S and 16S and considering the reading frame option for each CDS with default parameters (Katoh et al., 2017). The control region and tRNA were discarded as well as flanking regions that were not available for most terminals.

The final alignment consisted of 21,762 bp, comprising 14,115 bp for mtDNA (12S-16S: 2784; 11331 for mtDNA exons) and 7,647 bp for concatenated nuDNA (RAG1: 1374, NCX1: 1278, POMC: 558, BDNF: 693, CXCR4: 630, TYR: 675, SLC8A3: 1092, H3a: 288, RAG2: 744, RHO: 315). Two OTUs (*P. parva* and *P. carvalhoi*) had only partial mitogenomic data (12-16S, COI and Cytb and 16S and COI respectively) and 6 and 3 nuDNA loci available, respectively. All the other terminals had complete mitogenomes and at least 4 nuDNA loci (Supplementary table S3). Preliminary Maximum Likelihood analyses of each nuDNA loci using RAxML (see above for method) suggested overall topological congruence with mtDNA except the position of *P*. sp. "Guyana" (see results).

We selected the best-fit partition scheme and model of evolution for each partition using PartitionFinder V2.1.1 (Lanfear et al., 2017), according to the Bayesian Information Criterion (BIC) using the greedy scheme and linked branch length. We predefined 14 blocks, one for rRNA genes (12S and 16S), one for each codon position of concatenated mtDNA CDS regions, and one for each nuDNA CDS regions. This analysis resulted in a best partition scheme of seven partitions (1: 12-16S; 2: mtDNApos1, 3: mtDNApos2, 4: mtDNApos3, 5: POMC+RAG2+CXCR4+RAG1+TYR, 6: NCX1, 7: RHO, BDNF, SLC8A3, H3a).

We reconstructed a time-calibrated gene tree in BEAST 2.5.2 using a birth-death tree prior to account for extinction processes. We parameterized unlinked substitution models and unlinked clock models according to the models suggested by the PartitionFinder analysis. Trees were linked. Divergence time estimation was implemented using an uncorrelated relaxed lognormal clock model of the distribution of rates among branches for each partition (Drummond et al., 2006). We relied on secondary calibrations based on Hime et al. (2020), an extensive nuclear genomic dataset (220 loci 291 kb) of all major frog lineages and the last land bridge between South America and Africa ca. 105 Ma (Torsvik et al., 2008). We enforced the monophyly of Pipidae since this clade has been strongly supported in all phylogenetic (e.g. Irisarri et al., 2011, 2017; Cannatella, 2015; Feng et al., 2017; Hime et al., 2020) and paleontological analyses (e.g. Gomez, 2016). We also enforced the monophyly of African Pipidae (*Xenopus*, *Hymenochirus* and *Pseudhymenochirus*), thus favoring the topology in which *Pipa* is the sister group of other Pipidae following the results from the analyses of molecular data of Irissari et al. (2011, 2017); Feng et al. (2017); Hime et al. (2020) and of Cannatella (2015). This last work used a combination of molecular, morphological and fossil~~s resulting in~~ higher support for this topology and dates compatible with phylogenomic studies (Irissari et al., 2011, 2017; Feng et al., 2017; Hime et al., 2020). We acknowledge that the interrelationships among the three main Pipidae lineages remain contentious, but the scope of our study being the relationships and the timing of divergence within *Pipa* we believe these priors to be reasonable and should have no influence on the crown age of *Pipa* or any divergence times within the genus. Specifically, we assumed a uniform prior distribution for three nodes (1) the MRCA of Pipidae (between 129.1 the lower HPD from Hime et al., 2020 and 105.0 Ma the western Gondwana final break-up), (2) the crown age of African Pipidae (between 114.4 Ma and 85.9 Ma HPDs from Hime et al., 2020) and (3) the crown age of Pipoidae (between 172.6 and 150.0 Ma HPDs from Hime et al., 2020) which corresponds to the root of the tree.

We analyzed independently the four mtDNA partitions and the three nuDNA partitions because preliminary analyses suggested highly incongruent posterior distribution of node ages (regardless of the inclusion of 3rd codon position, the use of either Yule or Birth-Death, linkage of the clocks among mtDNA partitions and among nuDNA partitions i.e., two clock vs seven). The concatenated analyses ~~resulted in~~ intermediary age for the nodes among *Pipa* spp. that were thus considered incorrect. We set two independent Markov chain Monte Carlo (MCMC) runs of 200 million iterations each, recording every 10,000th iteration and using the first 10% of iterations as burn-in. We combined the log and the tree files and the resulting posterior samples of trees of the two independent runs using LogCombiner 2.5 (Bouckaert et al., 2014) and checked convergence of model parameters via time-series plots. Chain mixing was considered adequate when parameters achieved an effective sample size above 500 (obtained for all parameters). We extracted a maximum clade credibility tree (based on the 36,002 resulting trees) using Tree annotator 2.5 (Bouckaert et al., 2014).

2.4. Biogeographic analyses

We used the time-calibrated phylogeny obtained from nuDNA to infer ancestral areas and biogeographic events via the BioGeoBEARS package in R (Matzke, 2013). We compared three models: (i) a likelihood version of the Dispersal-Vicariance model (DIVALIKE; Ronquist, 1997) (ii) a likelihood version of the BayArea (BBM) model (Landis et al., 2013), and (iii) the Dispersal-Extinction Cladogenesis model (DEC; Ree & Smith, 2008). We also compared versions of these models allowing jump-dispersal as described by the J parameter (Matzke, 2013; Ree and Sanmartín, 2018; Klaus and Matzke, 2020). Models were compared using the Akaike Information Criterion (AIC).

To identify spatial processes of diversification, we considered five main geographic areas where known species currently occur: Guiana Shield (GS), Western Amazonia (WA), Brazilian Shield (BS), Atlantic Forest (AF) and a Trans-Andean region (TA). The three Amazonian regions correspond to major geological features of Amazonia (Hoorn et al., 2010) and to the large biogeographic regions known as Wallace’s districts (Wallace, 1854), roughly delimited by modern riverine barriers: the Madeira River, the Caquetá/Japurá – Solimões, and the lower course of the Amazon River. This spatial partitioning into three areas allows us to investigate the putative connectivity across Neotropical regions and dispersal routes across Amazonia (Réjaud et al., 2020; Fouquet et al., 2021). Even though we ~~only~~ included two populations of *Pipa snethlageae* (French Guiana and central Amazonia), mtDNA sequences are identical despite geographical distance. This large-sized species is associated with large swamps and in French Guiana, it reaches a northwestern distributional limit much like several other species associated with this habitat, such as *Leptodactylus intermedius*, *Hydrolaetare schmidti*, *Typhlonectes compressicauda*, *Dracaena guianensis*, *Paleosuchus niger,* among others (Lescure and Marty, 2000; Vacher et al., 2020; Gazoni et al., 2021) it is therefore likely that the range of this species extends throughout the Amazon basin and we thus considered its range to be panamazonian.

2.5. Teeth evolution and mode of reproduction

Several works have investigated the evolution of morphological and reproductive features within *Pipa* (Trueb and Cannatella, 1986, Canatella and Trueb, 1988, Trueb and Massemin, 2001), although not in the light of a robust molecular phylogeny and without including *P. arrabali* (see results, correspondence with the nominal *P. pipa* remained ~~also~~ ambiguous) and the candidate species identified herein (see results). We verified the presence (*P*. sp. "ES", *P*. sp. "Nordeste", *P*. sp. "South" and *P. arrabali*); or absence of teeth (*P. pipa*; *P*. sp. "WGU", *P*. sp. "WAM", P. sp. Negro", *P*. sp. "Central") in specimens corresponding to nine OTUs (see results) in addition to the ones already examined by Trueb and Cannatella (1986) and Trueb and Massemin (2001), under stereomicroscope LEICA MZ75.

Data on reproductive mode was available for all the outgroups and most candidate species (*P. aspera*: Trueb and Massemin, 2001), *P*. sp. "Guyana" (RE, pers. obs.), *P. arrabali* (Garda et al., 2006), *P. carvalhoi* (Fernandez et al., 2011), *P. myersi* (Trueb, 1984), *P. pipa* (Linnaeus, 1758); *P*. sp. "WGU" (RE, pers. obs.), *P. snethlageae* (Massemin et al., 2007), *P. parva* (Sokol, 1977). However, no data could be found for *P*. sp. "South", *P*. sp. "WAM", *P*.

sp. “Central”, *P*. sp. “Negro”, *P*. sp. “ES”, *P*. sp. “Nordeste” for which we assumed they display the same breeding behaviour ~~than~~ their close relatives

2.6. Rates shifts

We observed important discrepancies between the temporal estimates in the *Pipa* genus obtained from mtDNA vs. nuDNA, while the topologies are almost identical (see results). This discrepancy may be resolved by invoking two hypotheses: (1) nuDNA temporal estimates are correct and the mtDNA substitution rate is underestimated i.e. in fact greater in *Pipa* relative to the rest of the tree, or (2) mtDNA temporal estimates are correct and the nuDNA substitution rate is overestimated. In the absence of fossils it remains virtually impossible to tease these two hypotheses apart. However, given that the biogeographic history of codistributed frog groups is mostly circumscribed to the Neogene and that previous works have identified large variations in mtDNA rates across frog lineages (Irissari et al., 2012), we estimated that the first hypothesis is the most likely.

Consequently, we ~~carried out~~ a second phylogenetic analysis combining mtDNA and nuDNA, identical to our first analysis, but constraining the crown age of *Pipa* to the estimate obtained with nuDNA only (using the nuDNA 95% HPD interval as a uniform calibration prior). This additional analysis allowed us to estimate the average mtDNA rate within *Pipa*, conditional on a Neogene diversification of this genus, and to compare this rate with the rates found in other branches in the same tree, and with other studies ~~(e.g., in two~~ Neobatrachia groups *Allobates* (Réjaud et al., 2020) and *Boana* gr. *albopunctata* (Fouquet et al., 2021b~~)~~ for which mitogenomic data were analysed ~~in similar ways than~~ for *Pipa* (i.e., using the same partitioning and secondary time calibration from phylogenomic studies of Feng et al., 2017 and Hime et al., 2020). BEAST tree files were obtained from the authors. For each partition and phylogenetic analysis, we calculated the average rate applying to a specific subtree (e.g., in *Pipa*, in African pipids, etc.) as the average of the branch-specific rates (we used a relaxed molecular clock with one rate per branch) across all the branches of the subtree, weighted by the lengths of the branches.

2.7. Detecting changes in dN/dS

Simulation studies have shown that analyses of selection coefficients are rather robust to sequence divergence (Yang, 2006) (as is the case in the present study), having been successfully used in various studies with highly divergent species (e.g., Buschiazzo et al., 2012). In order to understand whether acceleration of evolutionary rates in *Pipa* is due to changes in selective pressure, we tested alternative models with different assumptions about ratios of non-synonymous/synonymous substitution rates (ω). The software PAML v.3.15 (Yang, 1997) was used to estimate the likelihood and the ω values of different models derived from the topologies and sequence information from single-gene alignments with all codon positions, as well as the mt and nuclear nucleotide data sets. Branch lengths were first optimized for each data set assuming a single ω for the whole tree, and they were fixed when all other parameters were estimated under alternative models. The null model had a single ω value for all branches, and it was compared against four alternatives, which allowed a second ω value on (i) the stem branch of *Pipa*, (ii) all *Pipa* branches. All models were compared using the AIC.

3. Results

3.1. mtDNA-based species delimitation

The phylogenetic trees obtained from the ML and the Bayesian analyses of the 16S locus strongly supported *Pipa* as monophyletic as well as the existence of four major clades within the genus. Two of these clades (*Pipa aspera*/*arrabali* and *Pipa pipa*/*snethlageae*) are mainly Amazonian (Figure 1A), one occurs in the Atlantic Forest (*Pipa carvalhoi*) and one in Trans-Andean regions (*Pipa parva*/*myersi*). Several deeply diverging lineages are also supported within *Pipa pipa*, *Pipa carvalhoi* and *Pipa arrabali*, indicating that the seven currently recognized species represent a vast underestimation of the actual species diversity within the genus (Figure 1A, S6). *Pipa arrabali* in particular is recovered paraphyletic with respect to *Pipa aspera*.

Of the three species~~-~~delimitation methods, mPTP was ~~found to be~~ the most conservative, delimiting 13 OTUs. With 14 delimited OTUs, ABGD ~~resulted in~~ a very similar partitioning. We kept the 12–17th partitions (P = 0.0092) based on two criteria: (i) they correspond to a plateau for group number, and (ii) it is close to the 1% arbitrary threshold of intraspecific divergence recognized in other vertebrate delimitation studies with the 16S locus (Puillandre et al., 2012). By contrast, GMYC delimited 21 OTUs (Figure 1A). Mean interspecific p distances among these OTUs reaches a minimum value of 1.6% between *P. aspera* and *P*. sp. “Guyana” and between *P. parva* and *P. myersi* and is below 3% in two more instances, between *P. arrabali* and *P*. sp. “South”, and between *P*. sp. “Negro” and *P*. sp. “Central” (Table S7). The consensus of the results obtained through the three methods led to the delimitation of 14 DNA-based OTUs (Figure 1A).

3.2. nuDNA differentiation

The relationships obtained using the four nuDNA loci are overall largely congruent with the mtDNA-based delimitation. For example, *Pipa* sp. “Central” is consistently recovered as sharing no alleles with the other three OTUs related to *P. pipa*. These three OTUs delimited within *Pipa pipa* are also recovered as sharing no alleles in NCX1, but this case is more ambiguous since allele sharing is observed on the other loci. The groups delimited within *P. carvalhoi* are also differentiated on RAG1 and NCX1. The case of *P. aspera* and the populations from Guyana (assigned to *P. arrabali*) is noteworthy since these species are recovered as a single OTU based on mtDNA. However, they do not share any alleles on any nuDNA loci except RAG1. Considering their distinct morphology and their divergence on nuDNA loci we considered them as distinct, leading to 15 OTUs.

3.3. Molecular dating

We assigned the best-fit models suggested by the PartitionFinder analysis to each of the seven partitions. The two combined BEAST analyses of both the mtDNA and the nuDNA data led to all parameters having ESS > 2000. The resulting phylogenetic relationships are all highly supported in both the mtDNA and the nuDNA trees (posterior probability pp > 0.99) and completely congruent between the mtDNA and the nuDNA, except the position of *Pipa* sp. "Guyana", which is supported as the sister species of *P. aspera* with mtDNA and as the sister species of the clade formed by *P. arrabali* and *P.* sp. "South" according to nuDNA. *Pipa* is structured in two major clades, one centered in the eastern part of its range (*P. carvalhoi* - Atlantic Forest; *P. aspera*, *P. arrabali* - Eastern Amazonia), and the other clade in the western part (*P. myersi*, *P. parva* - Trans-Andes and *P. pipa*, *P. snethlageae* - Amazonia; Figure 2).

However, major differences are observed in the timing of divergence between the two trees. The crown age of *Pipa* is found to date back to ca. 85 Ma (76–94, 95% HPD) according to mtDNA (Figure 2A), but to ca. 18 Ma (14–22, 95% HPD), i.e., 4.7 times younger according to nuDNA (Figure 2B). Since these dates are largely incompatible, either the rates of the mtDNA, the nuDNA, or both are erroneously estimated. In the absence of fossils branching along the *Pipa* stem, this question will remain contentious. Nevertheless, the mtDNA rates are notoriously more subject to underestimation by BEAST notably for ancient periods (Molak and Ho, 2015) as ~~it~~ could be the case for *Pipa*. Therefore, we considered an additional analysis based on the crown age obtained from the analysis of nuDNA (see below) and favoured the relationships and temporal estimates of the nuDNA analysis hereafter.

3.4. Rates shifts

The posterior distributions of the rates of each mtDNA partition analysed alone and with the three old priors (Cretaceous) is relatively stable across the tree i.e., among African pipids and for the stem and the crown for *Pipa* (Figure 3). The only notable exception is the high rate of the stem of *Pipa* for pos2 using Cretaceous calibrations compared to the ones of other branches, while this rate is similar in other partitions. Moreover, when adding the Neogene prior on the crown age of *Pipa* (based on the results with nuDNA), the rates were found to be, for all partitions, about four times higher in crown *Pipa*, and half as high in the stem, as compared to the analysis without the Neogene prior. The rate in African pipids was similar in both analyses. The rates estimated in *Allobates* and *Boana* are comparable (although slightly higher) to the rates we estimated in crown Pipa with the Neogene prior, except for pos 3 (Figure 3). The rates of the nuDNA partitions were similar across branches (Supplementary figure S9).

3.5. Changes in dN/dS

We compared a model with ω constant across the whole tree (single-ω model) with a "three-ω model" assuming independent ω values in (1) the stem branch of *Pipa*, (2) the branches within the crown of *Pipa* and (3) in other branches of the tree. The ω value estimated with the single-ω model was well below 1 for all different mt and nuclear genes (0.012–0.194) (Table 1), indicating the action of purifying selection to maintain gene function (Castallena et al. 2011).

With the three-ω model, for all individual mt genes (except *atp8* but consists of a short 159 bp locus), ω was always estimated to be higher (ca. X2) in the stem branch of *Pipa* than within crown *Pipa*, or in the other Pipoidea branches (Table 1). The ω estimate for stem *Pipa* was also about twice higher than the overall rate estimated with the single-ω model (Table 1). The three-ω model significantly outperformed (ΔAIC < 2) the single-ω model for seven genes (*cob*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*), and the single-ω model significantly outperformed the three-ω model for two genes (*atp8* and *cox1*), as shown in Table 1. When all mitochondrial genes were considered jointly, the three-ω model was favoured overwhelmingly (ΔAIC = 124, Table 1). For nuclear genes (*h3a* having been excluded given this locus was short - 288 bp - and only represented by six terminals including a single *Pipa*), the three-ω model never significantly outperformed (ΔAIC < 2) the single-ω model, and the single-ω model was best for four genes (*bdnf*, *rag2*, *rho* and *slc8a3*), see Table 1.

Overall, our results suggest a relaxation of the purifying selection acting on mtDNA in *Pipa* which could be responsible, at least in part, for the general acceleration of the mtDNA substitution rates in *Pipa*. It is important to note that results derived from the comparison of different selection regimes should be taken with caution, because the analyzed sequences are highly divergent and silent substitutions might be saturated, thus compromising the correct estimation of ω values (Yang 2006).

3.6. Biogeography, reproductive mode evolution and teeth loss

Pipa diversified throughout the Neogene with two concomitant splits ca. 13 Ma between Amazonia and the Atlantic Forest and between Amazonia and Trans-Andes. Subsequent diversification within *Pipa* is circumscribed to the Pliocene and Pleistocene i.e., <5 Ma. Inference of ancestral areas using BioGeoBEARS equally favored the DEC and DEC+J models (Table S8) and provided very ambiguous results for the early diversification events within *Pipa*. Nevertheless, when the three amazonian areas are ~~considered~~ combined an amazonian origin is supported for the genus and for the two major clades. Concerning the clade (*P. pipa/P. snethlageae*) the combined probabilities suggest ~~rather~~ a western Amazonian origin and a recent (Middle Pleistocene) dispersal to the Guiana Shield. In the opposite direction but concomitantly, a dispersal from the Guiana Shield toward the Brazilian Shield is suggested for the clade (*P. aspera*/*P. arrabali*).

Endotrophy has either evolved independently in the two Amazonian clades (the “macropipa” and the *P. arrabali*/*P. aspera* clades) or has been acquired ancestrally and independently lost in the Atlantic Forest and the Trans-Andean clades. Maxillary teeth havebeen lost once during the early Miocene in the western clade formed by “macropipa” and the Trans-Andean clade

4. Discussion

Our results shed a new light on species diversity, phylogenetic relationships, divergence time, biogeography and the evolution of edentulism and of reproductive mode in the genus *Pipa*. Our comparative results on mt and nuDNA also strikingly exemplify how molecular dating based on single locus (even though full mitogenomes) could lead to spurious time estimates when substitution rates vary drastically along long branches and no fossils are available along these long branches.

4.1. Species diversity

Almost all of the studies that have explored the question of how many species exist within particular amphibian clades in the Neotropics have uncovered high numbers of candidate species. Most often these cases correspond to populations that were previously considered to belong to species with wide distributions (e.g., Fouquet et al., 2014; 2021a,b; Gehara et al., 2014; Carvalho et al., 2021; Vacher et al., 2020). This situation is also exemplified in our results for *Pipa*, with a possible 2.1 times increase in species diversity (15 instead of seven) since eight OTUs possibly correspond to unnamed species (or taxa requiring revalidation such as *Pipa laevis* Cuvier, 1831 from the Rio Negro), in addition to the seven species already described. This figure is not surprising and even lower than what has been found in many other codistributed clades. Among these nine candidate species, most are supported as distinct by nuDNA; however, the situation remains ambiguous between some closely related pairs such as *P. arrabali* vs. *P.* sp “South”; *P. pipa* vs. *P.* sp. “WGU” and *P.* sp. “WAM”; *P. carvalhoi* vs *P.* sp. “Nordeste” and *P.* sp. “Central” vs *P.* sp. “Negro”, i.e., could represent false positives, either because allele sharing suggests conspecificity or simply because of the absence of data. The cases of *P. carvalhoi*, *P.* sp. “ES” and *P.* sp. “Nordeste” were already documented by Lima et al. (2020), who found little morphological, but clear chromosomal differences. Interestingly, interspecific genetic distances on 16S among closely related OTUs are lower than 3% in four instances (Table S7), notably between *P. aspera* and *P.* sp. “Guyana” (1.6%) and between *P. parva* and *P. myersi* (1.6%). This 3% threshold in 16S rDNA distances has frequently been suggested to be indicative of candidate species in anurans (Fouquet et al., 2007; Vieites et al., 2009). Nonetheless, these two pairs are phenotypically distinct (Trueb, 1984; Trueb and Cannatella, 1986), and congruent nuDNA divergences are recovered, thus confirming their status as distinct species. *Pipa aspera* and *P.* sp. “Guyana” illustrate a case of a false negative since mtDNA-based species delimitation actually failed to distinguish them. ~~The use of single genetic marker distance criteria may inherently result in a proportion of false negatives and false positives using short mtDNA locus and should only be used as a preliminary approximation of species boundaries and with the necessary precaution in species delimitation.~~ Here more data are needed, notably spatial sampling, morphology and bioacoustics, and we once more advocate the pertinence of integrative approaches that combine both genetic and phenotypic data (Padial et al., 2010).

Our results also highlight that the recognized species’ geographic ranges of *Pipa pipa* (possibly limited to Eastern Amazonia), *P. arrabali* (possibly limited to the northern part of the Madeira-Tapajos interfluvium) and *P. carvalhoi* (possibly limited to the Brazilian states of Bahia, Pernambuco and Alagoas) may be much more restricted than currently ~~admit~~ted. Moreover, it is likely that additional species remain unsampled, notably in southeastern Amazonia where records of *P. arrabali* not included in this study ~~have been documented~~ (e.g., Garda et al., 2006 from Serra do Cachimbo, Para; da Silva et al., 2020 in Tocantins; Pinheiro et al., 2012 from Carajas, Para), and in the Orinoco basin where two species of “macropipa” not included here may occur (Acosta-Galvis et al., 2016). The extent of these basic knowledge gaps for such a charismatic group of frogs is particularly striking and exemplifies the logistical challenges associated with undertaking fieldwork in remote Amazonian regions. Nevertheless, even considering 15 candidate species, *Pipa* remains a relatively poorly diversified clade as compared to most codistributed clades of terrestrial frogs (e.g., Réjaud et al., 2020) that also diversified throughout the Neogene.

4.2. Acceleration of mtDNA substitution rate in Pipa

Due to their small size, lack of recombination, rapid evolution rate, conserved gene content and genomic organization, and maternal inheritance, mitogenomes have been commonly used for analyzing phylogenetic relationships and divergence time (Igawa et al., 2008; Zhang et al., 2021). Although mitogenomes are usually thought to be under [illegible]tral or nearly neutral selection, evidence has accumulated for positive selection associated with environmental adaptations acting on mitochondrial genes (Carapelli et al., 2019; Shen et al., 2010). When shifts in selective forces occur in groups with long branches, along which no other constraints are available, this could lead to important bias in estimating these rates and divergence times.

Here we hypothesize that an acceleration of the substitution rate of mtDNA in *Pipa* relative to other pipoids largely explains incongruences across temporal estimates. This scenario is probable since (1) [illegible]iation of ω values is detected on all mtDNA genes, not only on a few genes which could have indicated positive selection on particular mt loci; (2) biogeography of codistributed group such as *Allobates*, *Amazophrynella*, *Adenomera*, *Pristimantis*, *Chiasmocleis* for instance display many topological and temporal similarities with *Pipa* when considering the nuDNA analysis, implying a mtDNA rate ca. four times higher than in other pipids, which still remains lower than in Neobatrachia; (3) Irisarri et al. (2012) already observed this pattern when comparing mt and nuDNA molecular evolution in anurans.

Irissari et al. (2012) hypothesised that relaxed purifying selection on mtDNA may at least partly explain the acceleration of mtDNA rates observed in Neobatrachia and discussed a number of possible reasons for such changes, notably in life history traits. In *Pipa*, the acquisition of skin incubation of the eggs and/or endotrophic development in the two amazonian clades could be related to that acceleration of substitution rates, although the precise causality remains totally speculative. Nevertheless, relaxed purifying selection on mtDNA has been hypothesized in other groups, such as in salamanders, concomitantly with the acquisition of direct development (Kakehashi and Kurabayashi, 2021). However, a somewhat more trivial, but not mutually exclusive, hypothesis may explain the observed variation in dN/dS ratio: the synonymous substitution saturation effect (i.e., several substitutions occurring sequentially at the same site on the same branch of the phylogenetic tree). This bias may particularly affect dN/dS estimation on very long branches, since dN/dS increases with saturation (Cannarozzi and Schneider, 2012). In other words, saturation leads to underestimate dS for ancient time periods because saturation is reached faster on sites that have higher substitution rates (e.g., pos1 and pos3) and thus the relative amount of dN (e.g., pos2) is artificially overestimated. This phenomenon is actually observable on Figure 3, in which the rate of the stem of *Pipa* of pos2 using Cretaceous calibrations is the highest. Consequently, the analysis via BEAST cannot correctly estimate mtDNA substitutions rates on the long stem of *Pipa* and overestimates the divergence time. This may also have been the case in Neobatrachia (Irissarri et al., 2012) and relaxed purifying selection may in fact not be the main process explaining the acceleration of mtDNA rates. Ancient divergence times (50–100 Ma) may simply be too old to be estimated using mitogenomes of lineages with rapid substitution rates such as *Pipa*, in the absence of fossil constraints from the Tertiary allowing substitution rates to be bounded.

4.3. Phylogeographic patterns in Amazonian Pipa

Within the *Pipa pipa/snethlageae* clade, the biogeographic analysis suggests a western Amazonian origin that is driven by ~~the fact that~~ early diverging lineages ~~occur~~ in western

Amazonia and in the Negro River. This western origin is expected given ~~the ecology of~~ these species tightly linked to large bodies of water that are currently widespread in the seasonal floodplains of the region~~, contrary to~~ the Guiana and the Brazilian Shields. Moreover, the historic presence of large lacustrine systems (Acre and Pebas systems; see Hoorn et al., 2010) in western Amazonia and the increasing eastward expansion of the flooded ecosystems since late Pleistocene (Aleixo and Rossetti, 2007, Bicudo et al., 2019) also support that hypothesis. However, the existence of five OTUs recovered within what has been so far considered ~~as~~ *P. pipa* is more surprising since it suggests limited connectivity across main amazonian tributaries whereas one would expect these large aquatic frogs to ~~efficiently~~ disperse along the vast hydrological network of the Amazon basin. In fact, there is an apparent paradox in *P. pipa* because it does display a shallow genetic structure throughout eastern Amazonia (Purus, Madeira, Tapajos, Xingu), but also including the easternmost part of the Guiana Shield (from Amapa, Brazil, to a western limit in the Suriname River in Suriname), while distinct lineages occupy western Amazonia (*P*. sp. “WAM”) and the Negro River (*P*. sp. “Negro” and *P*. sp. “Central”). The rivers of the Guiana Shield are not directly connected to the Amazon basin (although recent connection through the Rio Branco existed; de Souza et al., 2020) but the rivers of western Amazonia and the Negro River are. This pattern, along with the existence of two diverging lineages along the Negro river, is particularly noteworthy since spatially similar genetic structures are observed in fish (Piranhas: Hubert et al., 2007), in *Chelus fimbriata* (Vargas-Ramirez et al., 2020), ~~with~~ which *Pipa pipa* may actually be ecologically closer than to other frogs; but also in birds associated with flooded forests (Thom et al., 2020). Therefore, we can hypothesize that this pattern originates from either historical or ecological processes. Avulsion between the Japura and the course of the lower Negro river during Pleistocene (Ruokolainen et al., 2018) may explain this pattern; i.e., with ancient barriers that no longer exist; but also major paleo-mega-wetlands that were scattered in Amazonia during late

Neogene (Albert et al., 2018). Currently the strong ecological differences between the main types of floodplains (white waters in western Amazonia vs. black waters of the Negro, Cooke et al., 2014; Beheregaray et al., 2015; Oliveira et al., 2019) may also play a role in maintaining ecological isolation of the Rio Negro populations from the rest of the Amazon basin. Nevertheless, the nuDNA ~~only~~ support these difference for the Negro River OTUs~~ and caution should be taken given that the genetic structure observed on mtDNA may not be a reliable proxy of contemporaneous lack of interconnectivity among populations~~. Therefore, more spatial and genomic data are needed to investigate at finer scale the phylogeography within *Pipa pipa*.

In contrast, the genetic structure in the *P. aspera/arrabali* clade was expected to be more pronounced than within *Pipa pipa* since these species are associated with smaller water bodies in terra firme forest of the Brazilian and the Guiana Shields. However, even if a marked genetic structure exists, divergences remain relatively low among these lineages and even across the Amazon River. Nevertheless, the distribution of the different OTUs is more similar to the one found in terrestrial anurans i.e., distribution breaks matching major Amazonian tributaries and bisecting the Guiana Shield (Vacher et al., 2020). Interestingly, we recovered a mtDNA/nuDNA topological incongruence in the *Pipa aspera/arrabali* clade with the Brazilian Shield OTUs being nested within Guiana Shield ones according to nuDNA vs. ~~forming two subclades~~ according to mtDNA. These topological differences suggest either (1) introgression with mtDNA capture among GS lineages after the dispersal of *P. arrabali* from the GS or (2) introgression between BS and Guyana populations after dispersal from the BS, or (3) different coalescent gene histories. The origin of the group in the Guiana Shield was supported by the biogeographic analysis and the lower diversity among Brazilian Shield populations on the different nuDNA loci supports the first hypothesis. However, both topologies imply trans-amazon dispersals, nuDNA suggesting relatively a recent event (<1.4

Ma) when the transcontinental Amazon river was already established which is surprising given the ecology of the species.

4.4. Biogeography

The geographic center of diversification of most Neotropical groups remains difficult to infer not only because of putative extinctions but also because subsequent dispersals and intense landscape dynamics reshuffled the spatial distribution of organisms throughout the Cenozoic (e.g. Antonelli et al., 2018). Consequently, our ability to investigate the spatial origins of focal groups is usually restricted to more recent periods (Smith et al., 2014; Marques Silva et al., 2019; Cracraft et al., 2020). ~~The case of~~ *Pipa* is no different, with little signal for the early ancestral range of the main groups. Nevertheless, since (1) two of the main *Pipa* lineages occupy Amazonia and that they possibly originated in the eastern (*P. aspera/P. arrabali*) and the western portions (*P. pipa/P. snethlageae*) of that region, and that (2) their respective sister groups occupies the Atlantic Forest (*P. carvalhoi*) on the east and the Trans-Andean region (*P. myersi/P. parva*) on the west, a likely scenario consists of an Amazonian origin (14–22 Ma) and subsequent dispersal/vicariance some 12–13 Ma towards neighbouring regions. Such a scenario seems particularly plausible since several co-occurring amphibians have similar trajectories with amazonian origin dating back to the early Neogene also displaying an east-west pattern in Amazonia as well as concomitant dispersal/vicariance with Atlantic Forest and Trans-Andean regions (see below).

The main split in *Pipa* separates an eastern Amazonia and the Atlantic Forest clade from a panamazonian (with a likely western Amazonian origin see above) and a Trans-Andean clade~~s~~. This East-West split dating back to early Neogene (14–22 Ma) coincides with the Pebas system, a large freshwater system initially connected to the Caribbean Sea, that may have covered a surface of up to 13% of current Amazonia (800,000 km²; Albert et al., 2018). The Pebas system supposedly formed during the early Miocene (23 Mya) and occupied most of the Western Amazonian lowlands until around 10–9 Ma, when this system was progressively drained eastward into the Atlantic Ocean and transitioned into the modern Amazon watershed (Albert et al., 2018; Hoorn et al., 2017). Recent advances in Amazonian biogeography of amphibians (Rojas et al., 2018; Réjaud et al. 2020; de Carvalho et al., 2021; Fouquet et al. 2021a,b) and also other groups of terrestrial vertebrates recovered this east-west pattern within Amazonia. Interestingly, the edantate species of *Pipa* correspond to this western clade which suggests a single loss, a much simpler scenario than what was hypothesised by Trueb and Cannatella (1986). The origin of edentulism in this group of *Pipa* remains highly speculative but may be linked with feeding habits in an ancestral habitat in a large lacustrine system. How these frogs actually feed has been investigated by Fernadez et al. (2017) and Cundall et al. (2017) but only in *P. pipa*. A comparison between edantate and dantate species may provide some insights about the evolution of feeding habits.

The Andes represent a formidable barrier to dispersal between Amazonia and the Trans-Andean region. The Andean orogeny has been an ongoing process over the last 40 Ma punctuated by several intensive phases notably 12 Ma (Mora et al., 2010; Hoorn et al., 2010). The 12–13 Ma dispersal vicariance between Amazonia and the Trans-Andean region in *Pipa* coincides with this middle Miocene phase of orogeny. Moreover, similar Trans-Andean events occurred in other forestrial groups such as *Allobates* (Réjaud et al., 2020), *Engystomops* (Funk et al., 2012) and several lineages of Dendrobatidae (Santos et al., 2009), *Dendropsophus* (Pirani et al., 2020) and *Pristimantis* (Mendoza et al., 2015). Conversely, in amphibian taxa with higher dispersal abilities, notably having frequently adapted to altitude, such as glassfrogs or toads, some species have been isolated by the Andes (Bessa-Silva et al., 2020; Castroviejo-Fisher et al., 2014), while other species even display a cross-Andean distribution (e.g. *Boana boans*, Caminer and Ron, 2020; *Cochranella resplendens*; Molina-Zuluaga et al., 2017). Given that *Pipa* spp. are exclusively found in lowlands, we hypothesize that the Andes acted as a barrier to dispersal 12 Ma onward since the intense orogeny phase.

Similarly, the savannas of the Cerrado that today separates Amazonia and the Atlantic Forest represent a barrier for most forest amphibians. Middle Miocene divergence appears relatively concomitant with those other Neotropical frog and lizard complexes occurring in both regions (*Adenomera*: Fouquet et al., 2014; *Adelophryne*: Fouquet et al., 2012a; *Dendropsophus*: Pirani et al., 2020; *Leposoma*: Pellegrino et al., 2011), as well as birds (Batalha-Filho et al., 2013). The occurrence of successive dispersal routes between Amazonia and the Atlantic Forest has been largely recognized since this pattern has been recovered for frogs and other vertebrates (Ledo and Colli, 2017). However, both timing and approximate location of these connections remain largely debated. Nevertheless, there is little doubt that recurrent phases of forest extension have allowed recurrent connectivity and exchanges of fauna between these forests notably during the Miocene. The Middle Miocene Climatic Optimum (MMCO) (17–15 Ma) was a warm and wet period followed by periods of stronger seasonality associated with open vegetation expansion (Steinthorsdottir et al., 2021). Such environmental changes may have isolated forest-adapted species. Batalha-Filho et al. (2013) hypothesized a southern pathway between these biomes during the middle Miocene (Por, 1992; Costa, 2003). Widespread vegetation opening is documented subsequent to the MMCO (Flower & Kennett, 1994) with drastic climatic changes linked to major uplift of the Andes (Hoorn et al., 2010) and sea-current modifications (Herold et al., 2009; Le Roux, 2011).

The *Pipa* species with endotrophic development, which correspond to the two amazonian clades, are thus paraphyletic. Furness and Capellini (2019) demonstrated that complex parental investment such as brooding and “viviparity” are very unlikely to undergo reversals. Therefore, assuming that skin incubation predates endotrophic development, this implies two independent acquisitions of endotrophy: one in macropipa and one in the eastern amazonian *P. aspera/arrabali* clade. ~~The fact that endotrophy was probably acquired twice only in Amazonia may not be a simple coincidence but implying any causality, such as more stable climatic conditions in Amazonia providing long suitable periods for incubation or stronger selective pressure from predators remains highly speculative. Moreover,~~ some reversals from endotrophy to exotrophy have been documented in *Anomalogossus* (Vacher et al., 2017), and in *Adenomera* (Fouquet et al., 2014) and since the reproductive mode of many species remain undocumented, the occurrence of endotrophy-exotrophy shifts may ~~have in fact~~ be more common than currently considered. Given that there are only two pairs of lineages with both modalities in *Pipa*, reversals from endotrophy to exotrophy in the Atlantic Forest and Trans-Andes clades remain thus plausible.

5. Conclusion

These findings shed new light on the diversity, genomic evolution, historical biogeography and the evolution of morphology and reproductive mode in *Pipa*. Much still remains to be investigated on each of these aspects, notably improving the sampling in Amazonia and combining morphology and bioacoustic with molecular data to delimit species and ~~describe~~ the actual diversity in the genus. Our understanding of the biogeography and the morphological evolution of *Pipa* also harbor large gaps that fossils from Amazonia (Antoine et al., 2021) and microtomography could help to fill. We hope that this work will foster these

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Permission to conduct biodiversity research in Guyana was provided by the EPA Guyana under research permit number 180609 BR 112, and fieldwork was made possible through the Iwokrama International Centre for Rain Forest Conservation and Development particularly R. Thomas and the forestry department of the Guyana Forestry Commission (GFCPRDD). We have gathered the material used in this study thanks to F. Starace, A. Snyder, T. Colston, L. Dupreez, O. Chaline, M. Dewynter, M. Teixeira Jr., R. Recoder, M. A. Sena, I. Prates, P. R. Melo Sampaio, V. Prates, S.M. de Souza, A. Camacho, J. M. Guellere, P. Dias, A. Réjaud, J. Chretien, V. Premel, M. Chouteau, C. Schneider, V. Pinheiro, E. Courtois, C. marty, L. Storti, F. Dal Vechio, I. Prates, P. Peloso, J. Gomes, M. Hölting. We are deeply indebted to Juliette Vallin, Quentin Martinez, Juliana Tanaka, Alexandre Réjaud, Uxue Suescun, Sophie Manzi, Sabrina Baroni, Manuel jr. who contributed providing access to material and with data production. We also warmly thank Fabien Condamine, Pierre-Olivier Antoine, David Blackburn, Jonathan Rolland and Iker Irrissari for sharing thoughts and ideas about our findings.

Funding

This study benefited from an “Investissement d’Avenir” grant managed by the Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01; TULIP, ref. ANR-10-LABX-0041; ANAEE-France: ANR-11-INBS-0001), French/Brazilian GUYAMAZON program action (IRD, CNRS, CTG, CIRAD and Brazilian Fundação de Amparo Pesquisa do Estado do

Amazonas-FAPEAM 062.00962/2018) and co-coordinated by A. Fouquet and F. P. Werneck. M. T. Rodrigues thanks the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo Pesquisa do Estado de São Paulo (FAPESP grant numbers: 2003/10335-8, 2011/50146-6 and NSF-FAPESP Dimensions of Biodiversity Program [grant numbers: BIOTA 2013/50297-0, NSF-DEB 1343578]) and the NASA. P. J. R. Kok thanks the Fonds voor Wetenschappelijk Onderzoek (FWO12A7614N and FWO12A7617N) for financial support. RE's work benefited from grants from the German Academic Exchange Service (DAAD) and Deutsche Forschungsgemeinschaft (DFG; ER 589/2-1). F. P. Werneck thanks CNPq (productivity fellowship: 311504/2020-5).

References

Acosta-Galvis, A.R., Lasso, C.A., Morales-Betancourt, M.A., 2016. Ranas del género *Pipa* (Anura: Pipidae) de la Orinoquia colombiana: nuevos registros y comentarios sobre su taxonomía, distribución e historia natural. Biota Colomb. 17, 105–116.

Albert, J.S., Val, P., Hoorn, C., 2018. The changing course of the Amazon River in the Neogene: center stage for Neotropical diversification. Neotrop. Ichthyol. 16(3), e180033.

Aleixo, A., de Fátima Rossetti, D., 2007. Avian gene trees, landscape evolution, and geology: towards a modern synthesis of Amazonian historical biogeography?. J. Ornithol. 148(2), 443–453.

Anisimova, M., Bielawski, J.P., Yang, Z., 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol. Biol. Evol. 18, 1585–1592.

Antoine, P.O., Yans, J., Castillo, A.A., Stutz, N., Abello, M.A., Adnet, S., Custódio, M.A., Benites-Palomino, A., Billet, G., Boivin, M., Herrera, F., 2021. Biotic community and landscape changes around the Eocene–Oligocene transition at Shapaja, Peruvian Amazonia: Regional or global drivers?. Glob. Planet. Change 202, p.103512.

Antonelli, A., Zizka, A., Carvalho, F.A., Scharn, R., Bacon, C.D., Silvestro, D., Condamine, F.L., 2018. Amazonia is the primary source of Neotropical biodiversity. PNAS 115(23), 6034–6039.

Báez, A.M., Scanferla, C.A., Agnolin, F.L., Cenizo, M., Reyes, M.D.L., 2008. Pipid frog from the Pleistocene of the Pampas of southern South America. J. Vertebr. Paleontol. 28(4), 1195–1198.

Bandelt, H., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16(1), 37–48.

Batalha-Filho, H., Fjeldså, J., Fabre, P.H., Miyaki, C.Y., 2013. Connections between the Atlantic and the Amazonian forest avifaunas represent distinct historical events. J. Ornithol. 154(1), 41–50.

Beheregaray, L.B., Cooke, G.M., Chao, N.L., Landguth, E.L., 2015. Ecological speciation in the tropics: insights from comparative genetic studies in Amazonia. Front. Genet. 5, 477.

Bessa-Silva, A., Vallinoto, M., Sampaio, I., Flores-Villela, O.A., Smith, E.N., Sequeira, F., 2020. The roles of vicariance and dispersal in the differentiation of two species of the *Rhinella marina* species complex. Mol. Phylogenet. Evol. 145, 106723.

Bewick, A.J., Chain, F.J., Heled, J., Evans, B.J., 2012. The pipid root. Syst. Biol. 61(6), 913–926.

Bicudo, T.C., Sacek, V., de Almeida, R.P., Bates, J.M., Ribas, C.C., 2019. Andean tectonics and Mantle Dynamics as a Pervasive Influence on Amazonian ecosystem. Sci. Rep. 9(1), 1–11.

Blanco-Torres, A., Báez, L., Patiño-Flores, E., Renjifo, J.M., 2013. Herpetofauna del valle medio del río Ranchería, La Guajira, Colombia. Rev. Biodivers. Neotrop. 3(2), 113–122.

Böhme, M., 2003. The Miocene climatic optimum: evidence from ectothermic vertebrates of Central Europe. Palaeogeogr. Palaeoclimatol. Palaeoecol. 195, 389–401.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C-H., Xie, D., Suchard, MA., Rambaut, A.,& Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10(4), e1003537.

Burki, F., Roger, A.J., Brown, M.W., Simpson, A.G., 2020. The new tree of eukaryotes. Trends Ecol Evol. 35(1), 43–55.

Buschiazzo, E., Ritland, C, Bohlmann, J., Ritland, K., 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. BMC Evol. Biol. 12, 8.

Caminer, M.A., Ron, S.R., 2020. Systematics of the *Boana semilineata* species group (Anura: Hylidae), with a description of two new species from Amazonian Ecuador. Zool. J. Linnean Soc. 190(1), 149–180.

Cannarozzi, G. M., Schneider, A,. 2012). Codon evolution: mechanisms and models. Oxford University Press.

Cannatella, D., 2015. *Xenopus* in space and time: fossils, node calibrations, tip-dating, and paleobiogeography. Cytogenet. Genome Res. 145(3–4), 283–301.

Cannatella, D.C., Trueb, L., 1988. Evolution of pipoid frogs: intergeneric relationships of the aquatic frog family Pipidae (Anura). Zool. J. Linnean Soc. 94(1), 1–38.

Carvalho, T.R., Moraes, L.J., Lima, A.P., Fouquet, A., Peloso, P.L., Pavan, D., Drummond, L., Rodrigues, M.T., Giaretta, A., Gordo, M., Neckel-Oliveira, S., Haddad, C.F., 2021. Systematics and historical biogeography of Neotropical foam-nesting frogs of the

Adenomera heyeri clade (Leptodactylidae), with the description of six new Amazonian species. Zool. J. Linnean Soc. 191(2), 395–433.

Castellana, S., Vicario, S., Saccone, C., 2011. Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. Genome Biol. Evol. 3, 1067–1079.

Castroviejo- Fisher, S., Guayasamin, J.M., Gonzalez- Voyer, A., Vilà, C., 2014. Neotropical diversification seen through glassfrogs. J. Biogeogr. 41(1), 66–80.

Cooke, G.M., Landguth, E.L., Beheregaray, L.B., 2014. Riverscape genetics identifies replicated ecological divergence across an Amazonian ecotone. Evolution 68(7), 1947–1960.

Cracraft, J., Ribas, C.C., d'Horta, F.M., Bates, J., Almeida, R.P., Aleixo, A., Boubli, J.P., Campbell, K.E., Cruz, F.W., Ferreira, M. and Fritz, S.C., Grohmann, C.H., Latrubesse, E.M., Lohmann, L.G., Musher, L.J., Nogueira, A., Sawakuchi, A.O., Baker, P., 2020. The origin and evolution of Amazonian species diversity, in: Rull, V., Carnaval A.C. (Eds.), Neotropical diversification: patterns and processes. Springer, Cham, pp. 225–244.

Cundall, D., Fernandez, E., Irish, F., 2017. The suction mechanism of the pipid frog, *Pipa pipa* (Linnaeus, 1758). J. Morphol. 278(9), 1229–1240.

Cuvier, G.L.C.F.D., 1831. The animal kingdom arranged in conformity with its organization. Translated from the French, with notes and additions by H. M'Murtrie. Volume 2. G. & C. & H. Carvill, New York.

da Fonte, L.F.M., Latombe, G., Gordo, M., Menin, M., de Almeida, A.P., Hui, C., Lötters, S., 2021. Amphibian diversity in the Amazonian floating meadows: a Hanski core-satellite species system. Ecography 44, 1–16.

Dal Vechio, F., Prates, I., Grazziotin, F.G., Zaher, H., Rodrigues, M.T., 2018. Phylogeography and historical demography of the arboreal pit viper *Bothrops bilineatus* (Serpentes, Crotalinae) reveal multiple connections between Amazonian and Atlantic rain forests. J. Biogeogr. 45(10), 2415–2426.

Dantas, S.P., Tavares, H.D., Pascoal, W., Saviato, M.J., Ávila, R.W., Vasconcelos, T.S., Oda, F.H., 2019. New distribution records from the Brazilian Cerrado and species distribution modelling of *Boana crepitans*, *Lithobates palmipes*, *Pipa pipa*, and *Micrurus h. hemprichii*. Biodiversity 20(4), 149–160.

Delfino, M., Sánchez-Villagra, M.R., 2018. A Late Miocene Pipine Frog from the Urumaco Formation, Venezuela. Ameghiniana 55(2), 210–214.

Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6(5), 361–375.

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

Donoghue, P.C., Yang, Z., 2016. The evolution of methods for establishing evolutionary timescales. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 371(1699), p.20160020.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Evans, B.J., Gansauge, M.T., Stanley, E.L., Furman, B.L., Cauret, C.M., Ofori-Boateng, C., Gvoždík, V., Streicher, J.W., Greenbaum, E., Tinsley, R.C., Meyer, M., 2019. *Xenopus fraseri*: Mr. Fraser, where did your frog come from?. PloS one 14(9), p.e0220892.

Ezard, T., Fujisawa, T., Barraclough, T.G., 2009. Splits: species' limits by threshold statistics. R package version 1, 29.

Feng, Y.J., Blackburn, D.C., Liang, D., Hillis, D.M., Wake, D.B., Cannatella, D.C., Zhang, P., 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. PNAS 114(29), E5864–E5870.

Fernandez, E., Irish, F., Cundall, D., 2017. How a frog, *Pipa pipa*, succeeds or fails in catching fish. Copeia 105(1), 108–119.

Fouquet, A., Loebmann, D., Castroviejo-Fisher, S., Padial, J.M., Orrico, V.G., Lyra, M.L., Roberto, I.J., Kok, P.J., Haddad, C.F., Rodrigues, M.T., 2012a. From Amazonia to the Atlantic forest: Molecular phylogeny of Phyzelaphryninae frogs reveals unexpected diversity and a striking biogeographic pattern emphasizing conservation challenges. Mol. Phylogenet. Evol. 65(2), 547–561.

Fouquet, A., Recoder, R., Teixeira Jr, M., Cassimiro, J., Amaro, R. C., Camacho, A., Damasceno, R., Carnaval, A.C., Moritz, C., Rodrigues, M.T., 2012b. Molecular phylogeny and morphometric analyses reveal deep divergence between Amazonia and Atlantic Forest species of *Dendrophryniscus*. Mol. Phylogenet. Evol. 62(3), 826–838.

Fouquet, A., Blotto, B.L., Maronna, M.M., Verdade, V.K., Juncá, F.A., de Sá, R., Rodrigues, M.T., 2013. Unexpected phylogenetic positions of the genera *Rupirana* and *Crossodactylodes* reveal insights into the biogeography and reproductive evolution of leptodactylid frogs. Mol. Phylogenet. Evol. 67(2), 445–457.

Fouquet, A., Cassini, C.S., Haddad, C.F.B., Pech, N., Rodrigues, M.T., 2014. Species delimitation, patterns of diversification and historical biogeography of the Neotropical frog genus *Adenomera* (Anura, Leptodactylidae). J. Biogeogr. 41(5), 855–870.

Fouquet A., Leblanc K., Framit M., Réjaud A., Rodrigues M.T., Castroviejo-Fisher S., Peloso P.L.V., Prates I., Manzi S., Suescun U., Baroni S., Moraes L.J.C.L., Recoder R., de Souza S.M., Dal Vecchio F., Camacho A., Guellere J.M., Rojas-Runjaic F.J.M., Gagliardi-Urrutia G., de Carvalho V.T., Gordo M., Menin M., Kok P.J.R, Hrbek T., Werneck F.P., Crawford A.J., Ron S.R., Mueses-Cisneros J.J., Rojas Zamora R.R., Pavan D., Simões P.I., Ernst R., Fabre A.C., 2021a. Species diversity and biogeography of an ancient frog clade from the Guiana Shield (Anura: Microhylidae: *Adelastes*, *Otophryne*, *Synapturanus*) exhibiting spectacular phenotypic diversification. Biol. J. Linn. Soc. 132(2), 233–256.

Fouquet, A., Marinho, P., Réjaud, A., Carvalho, T. R., Caminer, M. A., Jansen, M., Rainha, R.N., Rodrigues, M.T., Werneck, F.P., Lima, A.P., Hrbek, T., Giaretta, A.A., Venegas, P.J., Chavez, G. Ron, S., 2021b. Systematics and biogeography of the Boana albopunctata species group (Anura, Hylidae), with the description of two new species from Amazonia. Syst. Biodivers. 19(4), 375–399.

Funk, W.C., Caminer, M., Ron, S.R., 2012. High levels of cryptic species diversity uncovered in Amazonian frogs. Proc. R. Soc. B: Biol. Sci. 279(1734), 1806–1814.

Galvis, P.A., Mejía-Tobón, A., Rueda-Almonacid, J.V., 2011. Fauna silvestre de la reserva forestal protectora, Montes de Oca, La Guajira, Colombia. Corpoguajira, Riohacha, Colombia. pp. 822.

Garda, A.A., Biavati, G.M., Costa, G.C., 2006. Sexual dimorphism, female fertility, and diet of *Pipa arrabali* (Anura, Pipidae) in Serra do Cachimbo, Pará, Brazil. South American J. Herpetol. 1(1), 20–24.

Gazoni, T., Lyra, M.L., Ron, S.R., Strüssmann, C., Baldo, D., Narimatsu, H., Pansonato, A., Schneider, R.G., Giaretta, A.A., Haddad, C.F., Parise-Maltempi, P.P., 2021. Revisiting the systematics of the *Leptodactylus melanonotus* group (Anura: Leptodactylidae): Redescription of *L. petersii* and revalidation of its junior synonyms. Zool. Anz. 290, 117–134.

Gehara, M., Crawford, A.J., Orrico, V.G., Rodriguez, A., Loetters, S., Fouquet, A., Barrientos, L.S., Brusquetti, F., De la Riva, I., Ernst, R., Urrutia, G.G., 2014. High levels of diversity uncovered in a widespread nominal taxon: continental phylogeography of the Neotropical tree frog *Dendropsophus minutus*. PloS one 9(9), p.e103958.

Gissi, C., San Mauro, D., Pesole, G., Zardoya, R., 2006. Mitochondrial phylogeny of Anura (Amphibia): a case study of congruent phylogenetic reconstruction using amino acid and nucleotide characters. Gene 366,228–237.

Gómez, R.O., 2016. A new pipid frog from the Upper Cretaceous of Patagonia and early evolution of crown-group Pipidae. Cretac. Res. 62, 52–64.

Gómez, R.O., Pérez-Ben, C.M., 2019. Fossils reveal long-term continuous and parallel innovation in the sacro-caudo-pelvic complex of the highly aquatic pipid frogs. Front. Earth Sci. 7, 56.

Greven, H., 2011. Maternal adaptations to reproductive modes in amphibians, in: Norris, D.O., Lopez, E.M. (Eds.), Hormones and reproduction of vertebrates, Vol. 2. Academic Press pp. 117–141.

Hedtke, S.M., Morgan, M.J., Cannatella, D.C. and Hillis, D.M., 2013. Targeted enrichment: maximizing orthologous gene comparisons across deep evolutionary time. PloS one 8(7), p.e67908.

Hemmi, K., Kakehashi, R., Kambayashi, C., Du Preez, L., Minter, L., Furuno, N., Kurabayashi, A., 2020. Exceptional enlargement of the mitochondrial genome results from distinct causes in different rain frogs (Anura: Brevicipitidae: *Breviceps*). Int. J. Genomics, 2020.

Hime, P.M., Lemmon, A.R., Lemmon, E.C.M., Prendini, E., Brown, J.M., Thomson, R.C., Kratovil, J.D., Noonan, B.P., Pyron, R.A., Peloso, P.L., Kortyna, M.L., 2021.

Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. Syst. Biol. 70(1), 49–66.

Hoorn, C., Wesselingh, F.P., Ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Sarkinen, T.S.,Antonelli, A., 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330(6006), 927–931.

Hoorn, C., Bogotá-A, G. R., - 41 -Romero-Baez, M., Lammertsma, E. I., Flantua, S. G., Dantas, E. L., Dino, R., do Carmo, D.A., Chemale Jr, F., 2017. The Amazon at sea: Onset and stages of the Amazon River from a marine record, with special reference to Neogene plant turnover in the drainage basin. Glob. Planet. Change 153, 51–65.

Hubert, N., Duponchelle, F., Nunez, J., Garcia- Davila, C., Paugy, D., Renno, J.F., 2007. Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of the Neotropical ichthyofauna. Mol. Ecol. 16(10), 2115–2136.

Igawa, T., Kurabayashi, A., Usuki, C., Fujii, T., Sumida, M., 2008. Complete mitochondrial genomes of three neobatrachian anurans: a case study of divergence time estimation using different data and calibration settings. Gene 407(1–2), 116–129.

Irisarri, I., Baurain, D., Brinkmann, H., Delsuc, F., Sire, J.Y., Kupfer, A., Petersen, J., Jarek, M., Meyer, A., Vences, M., Philippe, H., 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nat. Ecol. Evol. 1(9), 1370–1378.

Irisarri, I., San Mauro, D., Abascal, F., Ohler, A., Vences, M., Zardoya, R., 2012. The origin of modern frogs (Neobatrachia) was accompanied by acceleration in mitochondrial and nuclear substitution rates. BMC Genomics 13(1), 1–19.

Irisarri, I., Vences, M., San Mauro, D., Glaw, F.,Zardoya, R., 2011. Reversal to air-driven sound production revealed by a molecular phylogeny of tongueless frogs, family Pipidae. BMC Evol. Biol. 11(1), 1–10.

Kakehashi, R., Kurabayashi, A., 2021. Patterns of natural selection on mitochondrial protein-coding genes in lungless salamanders: relaxed purifying selection and presence of positively selected codon sites in the family Plethodontidae. Int. J. Genomics, 2021.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., Flouri, T., 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics 33, 1630–1638.

Katoh, K., Rozewicki, J., Yamada, K.D., 2017. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinformatics 20, 1160–1166.

Klaus, K.V., Matzke, N.J., 2020. Statistical comparison of trait-dependent biogeographical models indicates that podocarpaceae dispersal is influenced by both seed cone traits and geographical distance. Syst. Biol. 69, 61–75.

Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mole. Biol. Evol. 34, 772–773.

Ledo, R.M.D., Colli, G.R., 2017. The historical connections between the Amazon and the Atlantic Forest revisited. J. Biogeogr. 44(11), 2551–2563.

Leigh, J.W., Bryant, D., 2015. PopART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6(9), 1110–1116.

Lescure, J., Marty, C., 2000. Atlas des amphibiens de Guyane. Collection patrimoines naturels. MNHN, Paris.

Lima, L.R., Bruschi, D.P., Do Nascimento, F.A.C., Scherrer De Araujo, P.V., Costa, L.P., Thomé, M.T.C., Garda, A.A., Zattera, M.L., Mott, T., 2020. Below the waterline: cryptic diversity of aquatic pipid frogs (*Pipa carvalhoi*) unveiled through an integrative taxonomy approach. Syst. Biodivers. 18(8), 771–783.

Linnaeus, C., 1758. Systema naturae (Vol. 1, No. part 1). Laurentii Salvii, Stockholm, p. 532.

Silva, S.M., Peterson, A.T., Carneiro, L., Burlamaqui, T.C.T., Ribas, C.C., Sousa-Neves, T., Miranda, L.S., Fernandes, A.M., d'Horta, F.M., Araújo-Silva, L.E., Batista, R., 2019. A dynamic continental moisture gradient drove Amazonian bird diversification. Sci. Adv. 5(7), p.eaat5752.

Massemin, D., Bordage, D., Kuntz, K., 2007. Report on the occurrence of *Pipa snethlageae* (Anura: Pipidae) in French Guiana, with notes on its natural history. Salamandra 43, 139–147.

Matzke, N.J., 2013. BioGeoBEARS: biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R Package, Version 0.2, 1, 2013.

Mendoza, Á.M., Ospina, O.E., Cárdenas-Henao, H., García-R, J.C., 2015. A likelihood inference of historical biogeography in the world's most diverse terrestrial vertebrate genus: Diversification of direct-developing frogs (Craugastoridae: *Pristimantis*) across the Neotropics. Mol. Phylogenet. Evol. 85, 50–58.

Mezzasalma, M., Glaw, F., Odierna, G., Petraccioli, A., Guarino, F.M., 2015. Karyological analyses of *Pseudhymenochirus merlini* and *Hymenochirus boettgeri* provide new insights into the chromosome evolution in the anuran family Pipidae. Zool. Anz. 258, 47–53.

Molak, M., Ho, S.Y., 2015. Prolonged decay of molecular rate estimates for metazoan mitochondrial DNA. PeerJ 3, e821.

Molina-Zuluaga, C., Cano, E., Restrepo, A., Rada, M., Daza, J.M., 2017. Out of Amazonia: the unexpected trans-Andean distribution of *Cochranella resplendens* (Lynch and Duellman, 1978)(Anura: Centrolenidae). Zootaxa 4238(2), 268–274.

Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., Vogler, A.P., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst. Biol. 58, 298–311.

Mora, A., Baby, P., Roddaz, M., Parra, M., Brusset, S., Hermoza, W., Espurt, N., 2010. Tectonic history of the Andes and sub-Andean zones: implications for the development of the Amazon drainage basin, in: Hoorn, C. Wesselingh, F. (Eds.), Amazonia, landscape and species evolution: a look into the past. Wiley-Blackwell, pp. 38–60.

Motta, J., Menin, M., Almeida, A. P., Hrbek, T., Farias, I.P., 2018. When the unknown lives next door: a study of central Amazonian anurofauna. Zootaxa 4438(1), 79–104.

Oliveira, J.A., Farias, I.P., Costa, G C., Werneck, F.P., 2019. Model- based riverscape genetics: disentangling the roles of local and connectivity factors in shaping spatial genetic patterns of two Amazonian turtles with different dispersal abilities. Evol. Ecol. 33, 273–298.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. Front. Zool. 7(1), 1–14.

Pellegrino, K.C., Rodrigues, M.T., Harris, D.J., Yonenaga-Yassuda, Y., Sites Jr, J.W., 2011. Molecular phylogeny, biogeography and insights into the origin of parthenogenesis in the Neotropical genus *Leposoma* (Squamata: Gymnophthalmidae): Ancient links between the Atlantic Forest and Amazonia. Mol. Phylogenet. Evol. 61(2), 446–459.

Pinheiro, L.C., da Cunha Bitar, Y.O., Galatti, U., Neckel-Oliveira, S., dos Santos-Costa, M. C., 2012. Amphibians from southeastern state of Pará: Carajás Region, northern Brazil. Check List 8(4), 693–702.

Pirani, R.M., Peloso, P.L., Prado, J.R., Polo, É.M., Knowles, L.L., Ron, S.R., Rodrigues, M.T., Sturaro, M.J., Werneck, F. P., 2020. Diversification history of clown tree frogs in neotropical rainforests (Anura, Hylidae, *Dendropsophus leucophyllatus* group). Mol. Phylogenet. Evol. 150, 106877.

Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609.

Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877.

Prates, I., Penna, A., Rodrigues, M.T., Carnaval, A.C., 2018. Local adaptation in mainland anole lizards: Integrating population history and genome–environment associations. Ecol. Evol. 8(23), 11932–11944.

Rabb, G.B., Rabb, M.S., 1960. On the mating and egg-laying behavior of the Surinam toad, *Pipa pipa*. Copeia 1960(4), 271–276.

Ree, R.H., Sanmartín, I., 2018. Conceptual and statistical problems with the DEC +J model of founder-event speciation and its comparison with DEC via model selection. J. Biogeogr. 45, 741–749.

Ree, R.H., Smith, S.A., 2008. Maximum Likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.

Réjaud, A., Rodrigues, M.T., Crawford, A.J., Castroviejo- Fisher, S., Jaramillo, A.F., Chaparro, J.C., Glaw, F., Gagliardi- Urrutia, G., Moravec, J., De la Riva, I.J., Perez, P., 2020. Historical biogeography identifies a possible role of Miocene wetlands in the diversification of the Amazonian rocket frogs (Aromobatidae: *Allobates*). J. Biogeogr. 47(11), 2472–2482.

Rojas, R.R., Fouquet, A., Ron, S. R., Hernández-Ruz, E. J., Melo-Sampaio, P. R., Chaparro, J. C., Vogt, R.C., De Carvalho, V.T., Pinheiro, L.C., Avila, R.W., Farias, I., Gordo, M., Hrbek, T., 2018. A Pan-Amazonian species delimitation: high species diversity within the genus *Amazophrynella* (Anura: Bufonidae). PeerJ 6, e4941.

Ronquist, F., 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 195–203.

Ruokolainen, K., Moulatlet, G.M., Zuquim, G., Hoorn, C., Tuomisto, H., 2018. River network rearrangements in Amazonia shake biogeography and civil security. Preprints.

Santos, J.C., Coloma, L.A., Summers, K., Caldwell, J.P., Ree, R., Cannatella, D.C., 2009. Amazonian amphibian diversity is primarily derived from late Miocene Andean lineages. PLoS Biol. 7(3), e1000056.

Silva, L.A.D., Carvalho, P.S., Pereira, E.A., Fadel, R.M., Dantas, S.P., Brandão, R.A., Santana, D.J., 2020. Richness, diversity patterns, and taxonomic notes of amphibians from the Tocantins state. Biota Neotrop. 20(1).

Smith, B.T., McCormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Perez-Eman, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S., Brumfield, R.T., 2014. The drivers of tropical speciation. Nature 515(7527), 406–409.

Sokol, O.M., 1977. The free swimming *Pipa* larvae, with a review of pipid larvae and pipid phylogeny (Anura: Pipidae). J. Morphol. 154(3), 357–425.

de Souza, L.S., Armbruster, J.W., Willink, P.W., 2020. Connectivity of neotropical river basins in the central Guiana Shield based on fish distributions. Front. For. Glob. change 3, 8.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Steinthorsdottir, M., Coxall, H.K., De Boer, A.M., Huber, M., Barbolini, N., Bradshaw, C.D., Burls, N.J., Feakins, S.J., Gasson, E., Henderiks, J., Holbourn, A.E., 2021. The Miocene: The future of the past. Paleoceanogr. Paleoclimatology 36(4), p.e2020PA004037.

Streicher, J.W., Miller, E.C., Guerrero, P.C., Correa, C., Ortiz, J.C., Crawford, A.J., Pie, M.R., Wiens, J.J., 2018. Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. Mol. Phylogenet. Evol. 119, 128–143.

Thom, G., Xue, A. T., Sawakuchi, A. O., Ribas, C. C., Hickerson, M. J., Aleixo, A., Miyaki, C., 2020. Quaternary climate changes as speciation drivers in the Amazon floodplains. Sci. Adv. 6(11), eaax4718.

Torsvik, T.H., Müller, R.D., Van der Voo, R., Steinberger, B., Gaina, C., 2008. Global plate motion frames: toward a unified model. Rev. Geophys. 46(3), 1–44.

Trueb, L., 1984. Description of a new species of *Pipa* (Anura: Pipidae) from Panama. Herpetologica 40, 225–234.

Trueb, L., Cannatella, D.C., 1986. Systematics, morphology, and phylogeny of genus Pipa (Anura: Pipidae). Herpetologica 42, 412–449.

Trueb, L., Massemin, D., 2001. The osteology and relationships of Pipa aspera (Amphibia: Anura: Pipidae), with notes on its natural history in French Guiana. Amphibia-Reptilia *22*(1), 33–54.

Trueb, L., Ross, C.F., Smith, R., 2005. A new pipoid anuran from the Late Cretaceous of South Africa. J. Vertebr. Paleontol. 25(3), 533–547.

Vacher, J.P., 2017. Diversification in the Guiana Shield as seen through frogs (Doctoral dissertation, Université de Toulouse, Université Toulouse III-Paul Sabatier).

Vacher J.-P., Chave J., Ficetola F., Sommeria-Klein G., Tao S., Thébaud C., Blanc M., Camacho A., Cassimiro J., Colston T.J., Dewynter M., Ernst R., Gaucher P., Gomes J.O., Jairam R., Kok P.J.R., Dias Lima J., Martinez Q., Marty C., Noonan B.P., Nunes P.M.S., Ouboter P., Recoder R., Rodrigues M.T., Snyder A., de Souza S.M., Fouquet A. 2020. Large scale DNA-based survey of Amazonian frogs suggest a vast underestimation of species richness and endemism. J. Biogeogr. 47, 1781–1791.

Vargas-Ramírez, M., Caballero, S., Morales-Betancourt, M.A., Lasso, C.A., Amaya, L., Martínez, J.G., Viana, M.D.N.S., Vogt, R.C., Farias, I.P., Hrbek, T., Campbell, P.D., Fritz, U., 2020. Genomic analyses reveal two species of the matamata (Testudines: Chelidae: *Chelus* spp.) and clarify their phylogeography. Mol. Phylogenet. Evol. 148, 106823.

Wallace, A.R. 1854. On the monkeys of the Amazon. Ann. Mag. Nat. Hist. 14, 451–454.

Vaz-Silva, W., de Andrade, T.A., 2009. Amphibia, Anura, Pipidae, *Pipa pipa*: Distribution extension, new state record and geographic distribution map. Check List 5(3), 507-509.

Weygoldt, P., 1976. Beobachtungen zur biologie und ethologie von *Pipa* (*Hemipipa*) *carvalhoi* Miranda Ribeiro 1937 (Anura, Pipidae) 1. Zeitschrift für Tierpsychologie 40(1), 80–99.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.

Yang, Z., 2006. Computational molecular evolution. Oxford University Press, New York.

Zhang, J., Miao, G., Hu, S., Sun, Q., Ding, H., Ji, Z., Guo, P., Yan, S., Wang, C., Kan, X., Nie, L., 2021. Quantification and evolution of mitochondrial genome rearrangement in Amphibians. BMC Ecol. Evol. 21(1), 1–14.

Table 1

Results from the CodeML analyses for each locus and mitochondrial and nuclear loci combined. The highest ω values are in red and values in bold indicate models significantly outperforming the others according to the AIC.

Fig. 1. (A) Maximum Likelihood phylogenetic tree obtained from the analysis of 115 sequences of 16S (595 bp) of *Pipa*. Bootstraps > 50 % are indicated on the left side of the nodes and depicted with * when >99%. For sake of clarity, terminal branches are collapsed according to the results of the DNA-based species delimitation (ABGD, mPTP, GMYC) (complete tree given in Supplementary figure S6). Congruence across DNA-based species delimitation methods and with nuDNA neworks (absence of allele sharing) is illustrated by the use of colored columns on the right. Absence of available data for a given lineage is indicated with "NA". (B) Maps of northern South America showing the distribution of the sampled material color-coded as in the ML tree according to the species delimitation. (C) Median Joining networks based on four loci (NCX1 being divided in two, because parts of the individual sequences mostly overlapped either on 3' or the 5' ends) with corresponding colour code.

Fig. 2. Maximum clade credibility chronogram inferred in BEAST 2 based on (A) mitogenomic and (B) nuDNA, and (C) ancestral areas for *Pipa* inferred in BioGeoBEARS under the DEC model (results of the DEC+J model are available in Supplementary table S8). Nodes with maximum posterior probability are indicated with an asterisk under the branches. Calibrated nodes are indicated with a red circle. Node bars indicate the 95 % highest posterior distribution of node dates. Colored circles on the tips of the tree indicate the geographical distribution of sampled OTUs. Pie charts on nodes show the proportion of most likely

ancestral areas. Colors of node pie charts correspond to the geographic areas shown in the map. Changes in teeth, body size and larval development states are indicated with red bars.

Fig. 3. Comparison among posterior distributions, depicted as mustache plots, of the rates across the four mtDNA partitions either when mtDNA is analysed with calibrations dating back to the Cretaceous or when a Neogene calibration for the *Pipa* crown age is added (from nuDNA results). “Pipa” includes both “crown” and “stem”. Comparative rates for two groups of Neobatrachia (*Allobates* and *Boana* are also depicted for comparative purposes).

Figure 1 Click here to download Figure Figure 1.pdf

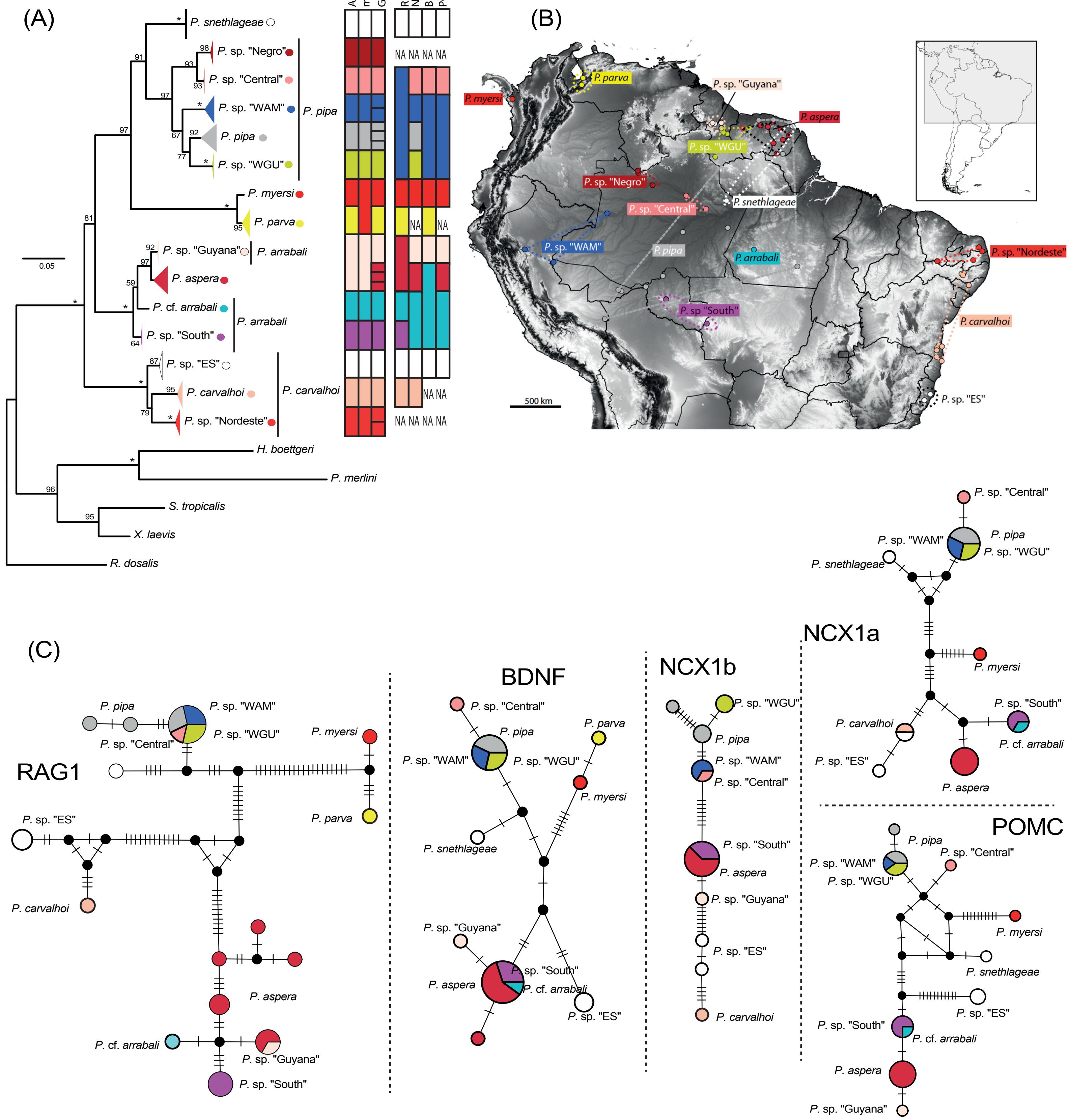

Figure 2

Click here to download Figure Figure 2.pdf

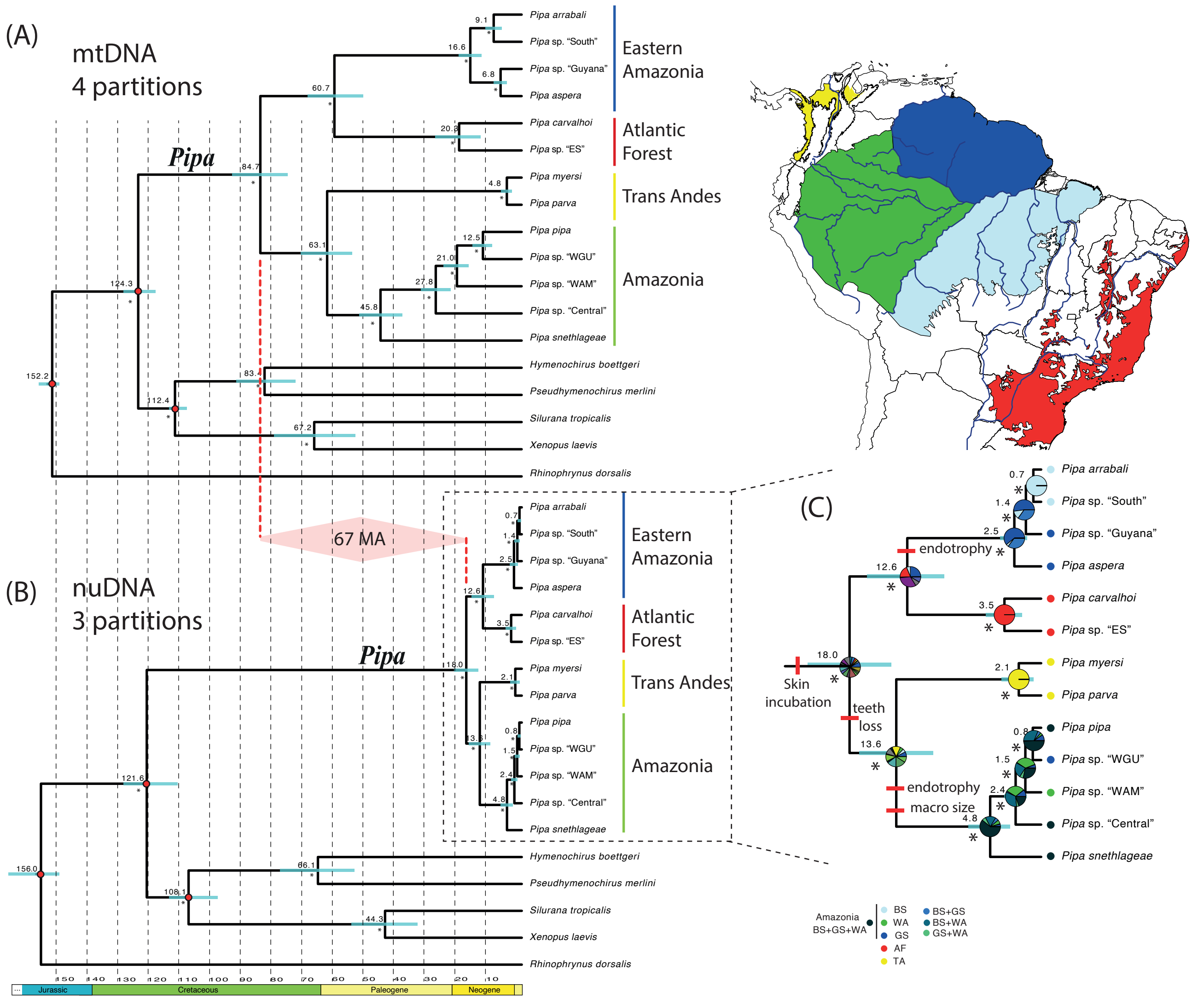

Figure 3

Click here to download Figure Figure 3.pdf

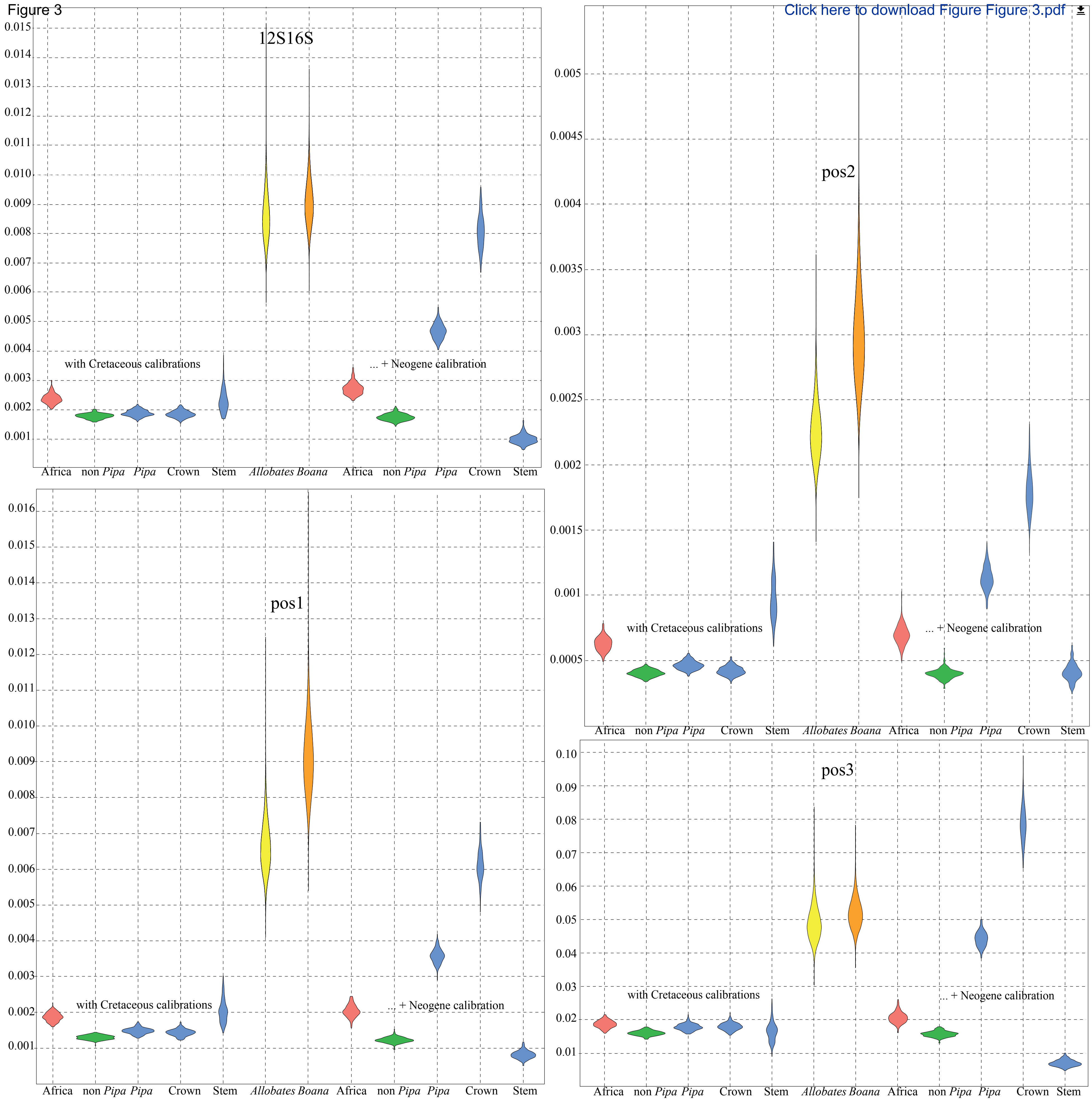

Table 1

Genes	Single-ω model (null)	Three-ω model			Delta AIC null	Delta AIC three-ω model
		non Pipa	Pipa crown	Pipa stem		
atp6	0.07458	0.06117	0.08291	0.13577	1.55	0.00
atp8	**0.16524**	0.16417	0.17029	0.12765	0.00	**3.79**
cob	0.05246	**0.05421**	**0.04754**	**0.16684**	**6.74**	0.00
cox1	**0.01312**	0.0131	0.0125	0.0272	0.00	**2.25**
cox2	0.03999	0.04029	0.03716	0.08989	0.00	0.81
cox3	0.03549	0.0322	0.03547	0.11395	0.78	0.00
nad1	0.07213	**0.083716**	**0.061665**	**0.168248**	**6.39**	0.00
nad2	0.1235	**0.14276**	**0.10034**	**0.2643**	**20.46**	0.00
nad3	0.1159	**0.14233**	**0.0899**	**0.21379**	**3.84**	0.00
nad4	0.09141	**0.08673**	**0.08854**	**0.21757**	**9.24**	0.00
nad4L	0.09169	**0.11485**	**0.06896**	**0.12563**	**2.43**	0.00
nad5	0.09683	**0.10923**	**0.08168**	**0.20561**	**23.26**	0.00
nad6	0.13259	0.15081	0.11618	0.15133	0.00	0.42
all mt prot-cod genes	0.08258	**0.09028**	**0.07151**	**0.19111**	**123.86**	0.00
bdnf	**0.0518**	0.05395	0.01992	0.06853	0.00	**2.63**
cxcr4	0.05564	0.04489	0.10037	0.07888	0.85	0.00
pomc	0.13991	0.15623	0.07654	0.10665	0.00	0.12
ncx1	0.04037	0.03954	0.06772	0.0263	0.00	1.35
rag1	0.08992	0.08365	0.12903	0.07478	0.52	0.00
rag2	**0.1939**	0.20566	0.13296	0.16237	0.00	**2.46**
rho	**0.09248**	0.0968	0.12761	0.04355	0.00	**2.50**
slc8a3	**0.03126**	0.02816	0.03978	0.04252	0.00	**3.11**
tyr	0.15588	0.12673	0.2416	0.18597	0.33	0.00
all nuclear genes	0.09206	0.08986	0.11794	0.08305	1.66	0.00