Original Article

# Geographic drivers of diversification in loliginid squids with an emphasis on the western Atlantic species

Gabrielle Genty[1]*, Carlos J Pardo-De la Hoz[1,2]*, Paola Montoya[1,3], Elena A. Ritschard[1,4]*

[1]Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá D.C, Colombia.

[2]Department of Biology, Duke University, Durham, North Carolina, 27708, United States of America

[3]Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, D.C., Colombia

[4]Department of Neuroscience and Developmental Biology, University of Vienna, Austria

* These authors contributed equally to this work.

Correspondence author: Gabrielle Genty, gy.genty180@uniandes.edu.co

## ABSTRACT

**Aim:** Identifying the mechanisms driving divergence in marine organisms is challenging as opportunities for allopatric isolation are less conspicuous than in terrestrial ecosystems. Here, we aim to estimate a dated phylogeny of the squid family Loliginidae, and perform ecological niche analyses to explore biogeographic and evolutionary patterns giving rise to extant lineages in this group, with particular focus on cryptic species with population structure along the western Atlantic coast.

**Location:** World-wide.

**Taxon:** Class Cephalopoda, Family Loliginidae

**Methods:** We used three loci to infer gene trees and perform species delimitation analysis to detect putative cryptic speciation events. We then estimated a dated species tree under the Bayesian multispecies coalescent and used it to reconstruct ancestral distributions based on the currently known ranges of the species. Also, we tested the hypothesis of niche divergence in three recently diverged species subpopulations of the northwestern and southwestern Atlantic Ocean by ecological niche modeling and niche overlap measurement from occurrence data.

**Results:** The phylogenetic analyses confirmed the monophyly for the current twenty-six species of the Loliginidae family. Our ancestral area reconstruction and divergence estimation revealed the origin and geographical dispersal of loliginid lineages. Additionally, the phylogenetic analysis and the species delimitation analysis supported geographic structure within *D. pleii, D. pealeii* and *L. brevis.* The ecological niche models revealed unsuitable habitat in the immediately adjacent area of the Amazonian Orinoco Plume, yet suitable habitat characteristics beyond this area.

**Main conclusions**: Our study allowed us to confirm the monophyly of all currently recognized species within the Loliginidae family and we corroborate the biogeographical origin being the Indo-Pacific region in the Cretaceous. We found a possible new cryptic lineage and show evidence of the Amazon-Orinoco Plume as an ecological barrier, which influenced the diversification of this particular group of marine organisms.

## Keywords

Loliginidae, Historical biogeography, ecological isolation, cryptic speciation

## INTRODUCTION

Geographic divergence can be seen as a continuum of processes causing isolation via dispersal barriers and ecological gradients varying across biogeographic regions (Cowman & Bellwood, 2013; Pyron & Burbrink, 2010). Hard physical barriers causing vicariance are often thought to be the major force driving divergence and the consequent speciation processes in terrestrial ecosystems (Coyne & Orr, 2004; Smith et al., 2014; Wiens, 2004). In contrast, the marine realm is considered to have fewer obvious physical barriers, which reduces the opportunities for strict allopatry to take place (Bowen et al., 2013; Brandley et al., 2010). Most marine invertebrates reproduce via a planktonic larval stage that is followed by either a benthic or a pelagic adult phase (Thorson, 1950). This life history trait confers them with the potential to facilitate extensive gene flow through dispersal across different geographic regions (Armonies, 2001; Cowen & Sponaugle, 2009; Simkanin et al., 2019). Yet, it has been estimated (Mora et al., 2011) that 2.2 million eukaryotic species exist in the oceans, and currently described marine species represent a large fraction of the global biodiversity (Costello et al., 2017; Foggo et al., 2003). The high estimated diversity suggests dynamic diversification processes beyond the apparent scarcity of geographic barriers in the ocean.

A number of mechanisms have been proposed to account for this astonishing biodiversity and current distributions of many groups of marine organisms. Habitat choice by settling larvae has been shown to be responsible for maintaining a mosaic genetic structure in two mussel species in a small geographic scale (Barrett, 2017; Bierne et al., 2003; Frolova & Miglietta, 2020); adaptation to alternative water temperatures maintains reproductive isolation between sister species of *Halichoeres* fishes (Rocha et al., 2005); reduced gene flow between continental and oceanic islands coasts has been shown to drive divergence in shallow-water reef-associated species (Hachich et al., 2015); sister lineages of Antarctic krill species have been shown to have diverged due to the formation of circum-Antarctic water circulation and the Antarctic Polar Frontal Zone (Patarnello et al., 1996); major dispersal barriers such as the Amazonian Orinoco Plume and the Mid-Atlantic Barrier, the Isthmus of Panama, among others, are also known to promote divergence as the result of vicariance or long distance dispersal across them, followed by local adaptation (Cowman & Bellwood, 2013; Luiz et al., 2013; Varona et al., 2019). Regardless of the spatial configuration in which the divergence takes place (*i.e.* allopatric, parapatric, peripatric or sympatric), the geographic and ecological heterogeneity plays a major role in all of these mechanisms.

Neocoleoid cephalopods are marine invertebrates that comprise octopuses, cuttlefishes and squids. Many studies have dealt with systematics and macroevolution in lineages of these organisms in the last decade (Kolis & Lieberman, 2019; Lindgren et al., 2012; Strugnell et al., 2006; Strugnell & Nishiguchi, 2007; Tanner, 2018; Wani, 2011). However, very few studies have addressed diversification dynamics from a geographic perspective, covering temporal and phylogenetic

hypotheses (but see (Amor et al., 2014; Brakoniecki, 1987; J. B. de L. Sales et al., 2017; Ulloa et al., 2017). Therefore, the geographic drivers of diversification in cephalopods remain poorly understood.

The increase of available molecular data for the species in the squid family Loliginidae, Lesueur, 1821, has resulted in systematic revisions with more phylogenetic accuracy. Evidence of possible cryptic species has been shown for many genera of loliginid squids(Anderson, 2000; J. B. de L. Sales et al., 2017), including three cases of potential cryptic speciation between lineages in the North and South Atlantic of the species Doryteuthis pealeii, D. pleii and Lolliguncula brevis (de Luna Sales et al., 2013; J. B. L. Sales et al., 2014). Also, a couple of studies have been published with divergence time estimates for major cephalopod lineages(Anderson, 2000; J. B. de L. Sales et al., 2017; Ulloa et al., 2017). Most species within this family are of commercial interest and their reproductive biology and ecology have been widely studied (Jereb & Roper, 2010; Martins et al., 2006; Okutani, 1995; Reid & Carstens, 2012; Vecchione, 1991). Due to their world-wide distribution (Fig. 1), and the availability of molecular and fossil data (M. Clarke & Maddock, 1988), loliginid squids represent an ideal case to study the mechanisms causing diversification in cephalopods.

**Figure 1.** Location of the worldwide distribution of loliginid squids with assigned distributions areas to each species according to previously defined marine biogeographic regions by Spalding et al. (2007).



We aim to use a data-mining approach to estimate a dated most-inclusive phylogeny of the family Loliginidae. We coupled this with ancestral area reconstruction and ecological niche modelling to explore patterns of diversification in a spatiotemporal context. This allowed us to assess the occurrence of potential diversification mechanisms driven by geographic factors. This includes (I)

5

isolation as a result of long-distance dispersal; (II) vicariance due to the presence of hard physical dispersal barriers; (III) divergence due to the presence of soft physical dispersal barriers (e.g. oceanic currents, river mouths); (IV) adaptation to different niches along ecological gradients.

## MATERIALS AND METHODS

### Taxon sampling and data acquisition

We used a data-mining approach to build our datasets using published sequences for 185 Operational Taxonomic Units (OTU's) comprising 26 loliginid species. We obtained 429 sequences for three markers (167 *Cytochrome Oxidase* I, 88 *Rhodopsin* and 126 *16S*) from the Genbank database (see Table 1 for references). Sequences from *Stenoteuthis oualensis* y *Omnastrephes bartramii* were used as outgroup in all analyses.

### Phylogenetic analysis and species delimitation

Multiple sequence alignments were performed using MAFFT 7(Katoh & Standley, 2013) with the Q-INS-i strategy as implemented in the web server (http://mafft.cbrc.jp/alignment/server/) and manually edited on MESQUITE 3.04 (Maddison & Maddison, 2007). PARTITIONFINDER 1.1.1 (Lanfear et al., 2012) was used to find the best partition scheme and model of nucleotide substitution for the single locus datasets using the rcluster search algorithm and the corrected Akaike Information Criterion (AICc). The searches were performed with the following starting subsets: *COI* $1_{st}$, $2_{nd}$, $3_{rd}$ codon positions; *RHO* $1_{st}$, $2_{nd}$, $3_{rd}$ codon positions and *16S*. To construct single locus trees, Maximum Likelihood analysis were done using RAXML 8 (Stamatakis, 2014) as implemented in the CIPRES online platform (Miller et al., 2010) using 1000 bootstrap and the GTR+CAT substitution model. Moreover, we generated single locus chronograms using an uncorrelated lognormal relaxed clock in BEAST v 1.8.3(Drummond et al., 2012; Drummond & Rambaut, 2007) with default priors. We ran two independent 10,000,000 generations analyses with four chains each sampling every $1000_{th}$ generation for a total sample size of 10,000 trees. Effective sample size for the estimated parameters and convergence was assessed in TRACER 1.6 (Andrew Rambaut et al., 2018). For each locus, a set of 400 trees from the estimated posterior distribution were subsampled using LOGCOMBINER 1.8.2 (A Rambaut & Drummond, 2015) after discarding 10% of the samples as burn-in. These sets of trees were finally used to infer single locus-based species delimitations using a bayesian implementation of the Generalized Mixed Yule Coalescent model (bGMYC) (Pons et al., 2006; Reid & Carstens, 2012).

### Species tree and divergence time estimation

Conflicting inter-species relationships were found with strong support (>70% bootstrap support) among phylogenies inferred from the different single-locus datasets. For this reason we decided to perform a dated species tree estimation in *BEAST (Heled & Drummond, 2010) as an alternative to concatenation. We assigned individuals to species according to the results of the previous phylogenetic analyses and used a yule tree prior for the inference. We used statoliths from the genus *Loligo* (M. R. Clarke & Fitch, 1979) to model a distribution of the node corresponding to the ancestor of the genus to calibrate the tree. For the calibration, we used a lognormal distribution

with the following parameters: log(mean)= 1.64; log(standard deviation)= 0.54; and Offset= 40. Ucld.mean parameter (Uncorrelated Lognormal relaxed clock mean) was set to a uniform prior distribution with an initial value of 0.002, a maximum value of 0.003 and a lower value of 0.001. Effective sample sizes for the estimated parameters and convergence were checked in TRACER 1.6. Consensus tree was generated on TREEANNOTATOR 1.8 (A Rambaut & Drummond, 2007) using a 50% majority rule and discarding 10% as burn-in.

**Ancestral areas reconstruction**

We used the Bayesian Binary MCMC and the Dispersal Extinction Cladogenesis (DEC) model as implemented in RASP (Ree & Smith, 2008; Yu et al., 2015) to obtain distribution ranges at each node. The inferred species tree was used as input for these analyses. We defined distributions areas (Fig. 1) according to previously defined marine biogeographic regions (Spalding et al., 2007) and assigned them to each species (Table S1).

**Climatic niche model approaches**

The Amazonian Orinoco Plume, in the southwest Atlantic Ocean is considered as a dispersal barrier for marine organisms (Luiz et al., 2013), since the decreases in the seawater salinity in this region makes it physiologically unsuitable for most non-estuarine marine organisms. Therefore, we tested whether the (a) Amazonian Orinoco Plume is functioning as a dispersal barrier for the putative species inside *D. pleii* and *L. brevis*, and/or (b) they are currently occupying different climatic spaces thus promoting divergence. No analysis was performed for *D. pealeii* because of the lack of geographic occurrences available at databases and/or papers. For this, we used ecological niche modelling and climatic space overlapping. In both cases, we obtained georeferenced occurrence data from GBIF database in June 2017 (Global Biodiversity Information Facility, www.gbif.org) and survey records in published literature (Herke & Foltz, 2002) for the species mentioned above. The data was handled at the population level (northern and southern populations), using only the environmental information available for its region of origin (i.e. North or South Atlantic). To estimate the climatic space and the potential geographic distribution for each, we used oceanic environmental layers of Salinity, pH, Sea Surface Temperature range, Chlorophyll-A range, Dissolved Oxygen, Diffuse Attenuation mean, and Photosynthetically Available Radiation from Bio-ORACLE (Tyberghein et al., 2012). Previous eco-physiological studies have shown that these variables affect the distribution of the modeled species (Laughlin & Livingston, 1982; Martins et al., 2006; Sobrino et al., 2002; Zielinski et al., 2000). We retrieved data from the climatic layers exclusively from the West Atlantic Ocean along America. We also restricted the layers to areas with up to 1000 m (Seibel, 2007) of depth using the bathymetrical information presented by Amante and Eakins (2009).

We generated separate climatic niche models for North and South populations using data from their region of origin, and then projected them from one region to the other. The ecological niche models were built using MAXENT 3.3.3 (S. Phillips et al., 2010; S. J. Phillips & Dudík, 2008), applying the default parameters and using cross-validation to evaluate the models with 20% data. We obtained logistic models to 1 km x 1 km, with values ranging from 0 (when the area is

unsuitable) to 1 (when is totally suitable) according to the climatic conditions. Finally, we used a multivariate environmental similarity surface (MESS) analysis (Elith et al., 2010) to evaluate the similarity between the North and South environments and therefore, the reliability of the projection.

We estimated the climatic space overlapping between northern and southern populations, following Broennimann and others (2012) using the *ecospat* package (Broennimann et al., 2015) as implemented in R Software (RStudio Team, 2020). The model performs a main components analysis using the environmental conditions of 10000 random locations throughout the entire study area (including both North and South regions). The climatic space estimation per population uses only the first two principal components (PC1 and PC2) and assumes a kernel distribution for the density of records to smoothen possible gaps. These gaps could be the result of sampling bias instead of conditions not being occupied by a population. The climatic space overlap is measured with the D metric (Schoener, 1970; Warren et al., 2017), ranging from 0 (when no overlap is found) to 1 (when the overlapping is complete). After that, a randomization process is used to evaluate if the estimated D value is significant according to the available conditions (Broennimann et al., 2012). A significantly low overlap between the climatic space of North and South populations could indicate climatic divergence between populations.

**RESULTS**

**Phylogeny and species delimitation**

The *COI* tree (Fig. 2) showed the highest resolution among the three genes, recovering the monophyly of the twenty-six loliginid species with a bootstrap >70% (except for *D. pleii*). Phylogenetic analyses based on the two other loci did not show strong support for all species, but they mostly recovered the same monophyletic groups or showed no significant conflict with the *COI* tree (Fig. S1, S2, S3). The phylogenetic analyses confirmed the previously reported geographic structure of *D. pleii, D. pealeii* and *L. brevis* (Anderson, 2000), with divergent populations in the North and South West Atlantic. This result was supported by the *COI* species delimitation analysis, in which the isolated populations resulted as distinct species (Fig. 2). Putative cryptic species or divergent lineages with geographic structure were detected inside three other clades (*Sepioteuthis lessoniana, Sepioteuthis sepioidea and Uroteuthis duvauceli*) with the Maximum Likelihood trees and the species delimitation analysis performed with *COI*.

We found incongruences regarding the relationship between species and monophyly of some genera. In order to solve the relationships between genera in the Loliginidae family, a phylogenetic analysis including all species with available sequences for the three genes was generated with a Bayesian coalescent approach (Fig. 3). This analysis clarified the evolutionary relationships between species and confirmed the monophyly from seven of the eight genera included in the study. The genus *Lolliguncula* is shown to be polyphyletic (Figure 2 and 3). This is supported by the species tree and the single locus phylogenies that show *L. panamensis* nested within the genus *Doryteuthis*.
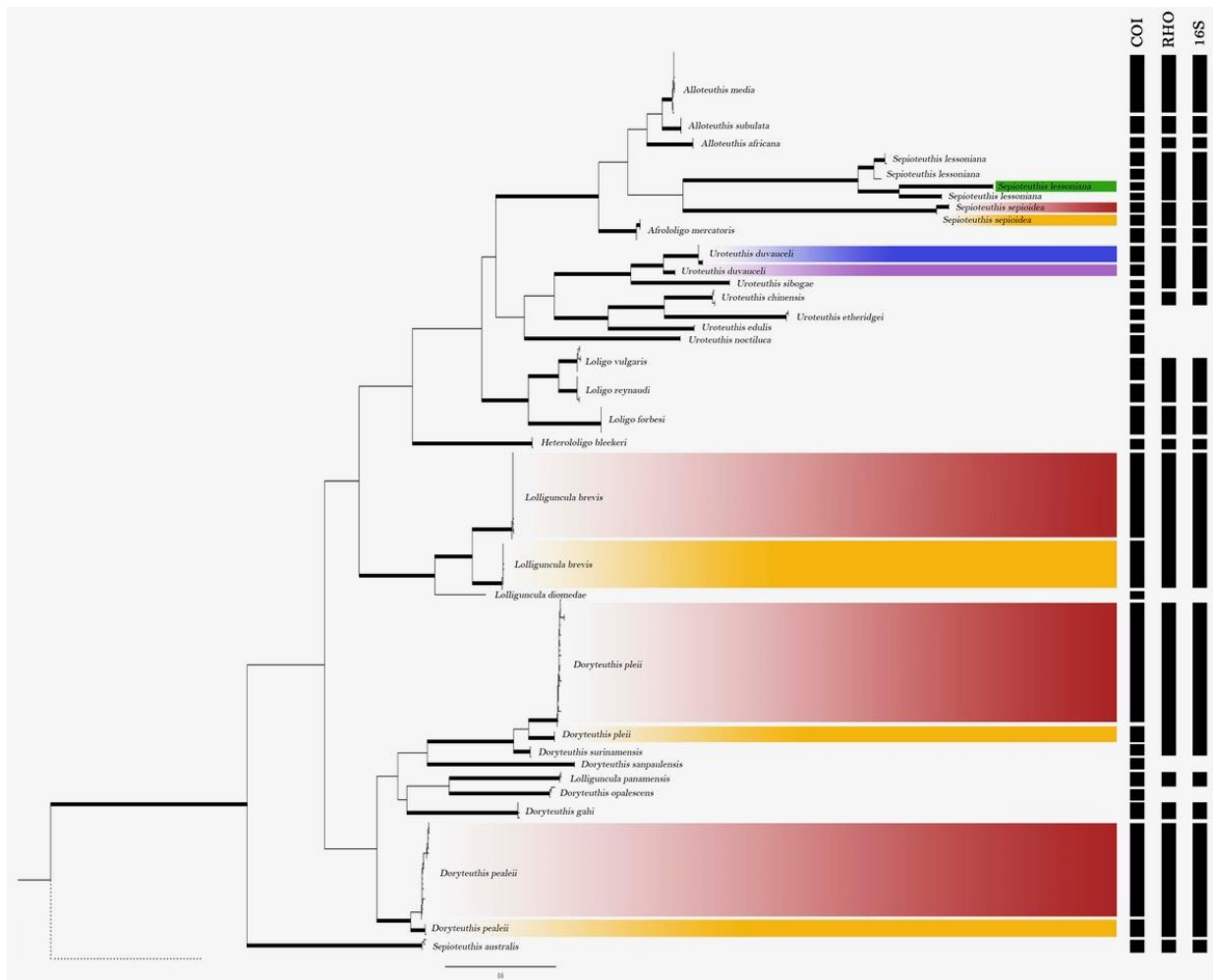
**Figure 2.** Maximum likelihood tree inferred from COI sequences. Thick branches represent bipartition with >70% bootstrap support. Bars to the right correspond to species delimitation inferred from bGMYC for each of the three loci used. Highlighted clades represent groups where cryptic speciation might have occurred and their current distribution as depicted in Figure 1.

### Ancestral area reconstruction and divergence estimation

The different models used to infer ancestral distribution ranges for nodes consistently estimated the origin and geographical dispersal of loliginid lineages (as shown in Fig. 3). A summary of the values estimated for dispersal, extinction, cladogenesis (in the DEC model), and dispersal and vicariance is presented in the Dryad Digital Repository and a description of the events is presented in the discussion.
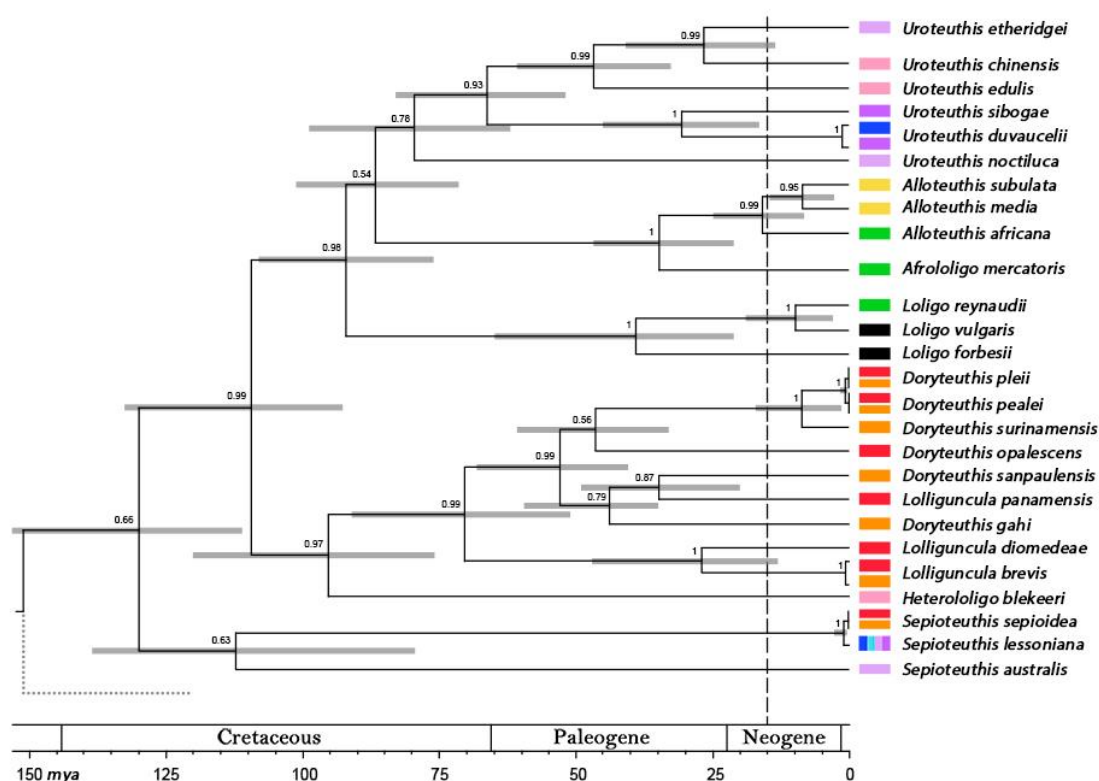
**Figure 3.** Dated species tree using information from three loci. Node bars represent 95% HPD of the node ages. Vertical discontinuous line with grey area indicated the estimated date of the closure of the Panamá isthmus. Color rectangles on the right show the current geographic distribution of each species as shown in Figure 1. Bottom scale shows geological time ages in mya.

**Climatic niche model approaches**

The ecological niche models (ENM) showed low probability of suitability in the immediately adjacent area of Amazonian Orinoco Plume for both species and increased as the areas are farther from the Plume (Fig. 4). According to the projections, the South region has suitable areas for North populations from both species (Fig. 4). In contrast, the North region has suitable areas only for the southern *L. brevis* (Fig. 4d). According to the MESS analyses the environmental conditions between the origin and projections sites are highly similar except in the South to North projection, where the northernmost region of the Atlantic present high uncertainty associated (Fig. S5).

In the climatic space overlap, we do not found evidence on climatic niche divergence between North and South populations for both species, since we found no significant overlap in the environmental space currently occupied between them (*L. brevis*: overlap=0.112, p-value=0.495; *D. pleii*: overlap=0.513, p-value=0.099).
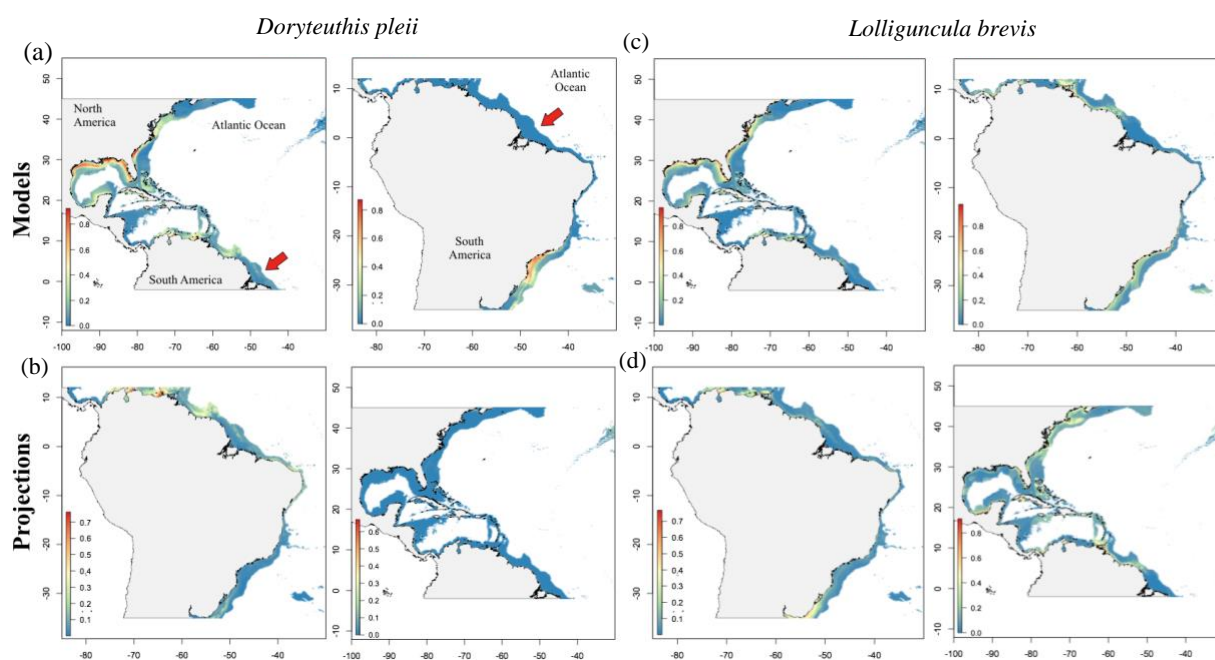
**Figure 4.** Ecological niche models for *D. pleii* and *L. brevis* and its projections. (a) and (c) show the potential geographic distribution for North and South populations of both species made by modelling the ecological niche. (b) and (d) show the projections of each model in the non-origin regions. Colors in the maps indicate the probability of suitability, with the highest probabilities in warm colors. The Amazonian Orinoco Plume region, marked with the red arrow, showed a low suitability for both species. According to the models, all populations have suitable areas in the non-origin regions, except for the South population of *D. pleii* in the North region as shown in (b).

## DISCUSSION

Having robust phylogenetic hypothesis is key to understand processes of diversification (Losos, 1996). Here we constructed a phylogeny of the family Loliginidae in order to unveil geographic drivers of diversification in these organisms. Together with ancestral areas reconstruction (AAR) and divergence time estimates among major clades, we aid the inference of the biogeographical history of the family in its early evolution. Conversely, species delimitation coupled with ecological niche model approaches revealed putative cryptic lineages that might represent recent and incipient diversification processes. The entire approach allowed us to correlate geological events and geographic heterogeneity with ancestral speciation and current distributions of extant lineages.

### New insights into phylogenetic relationships within Loliginidae

We confirmed the monophyly of all recognized species included here and clarified several between-species relationships within the family (Fig. 2 and 3). All phylogenetic analyses consistently showed the polyphyly of the genus *Lolliguncula,* as the species *L. panamensis* was nested within the *Doryteuthis* clade (Fig. 2, 3 S1, S2). The remaining two *Lolliguncula* species were monophyletic and appeared as the sister clade of the genus *Doryteuthis* (Fig. 2). The latest

phylogenetic revision of *Lolliguncula* showed *L. panamensis* as the earliest divergent lineage of the genus when using *Loligo* species as outgroup (Fig. 1 and 2 in (J. B. L. Sales et al., 2014), possibly resulted from a taxon sampling artifact as no sequences of others were included in the analyses. Our results suggest a review of *L. panamensis* and plausible re-circumscription for the species to *Doryteuthis panamensis*. Moreover, previous phylogenetic analyses using both mitochondrial and nuclear markers resulted in the polyphyly of the genus *Sepioteuthis* (de Luna Sales et al., 2013). We obtained the same topology using the *COI* dataset (Fig. 2), but the species tree recovered the three species from *Sepioteuthis* as monophyletic (Fig. 2). The inferred divergence time between *S. lessoniana* and *S. sepioidea* is very recent (Fig. 2), thus, incomplete lineage sorting due to deep coalescence might explain the conflicting results found between species and gene trees. Lastly, the position of *Heterololigo bleekeri* (Ulloa et al., 2017) is confirmed with our analyses, placing this species as a sister group to the *Doryteuthis/Lolliguncula* clade with strong support (Fig. 2).

**Migration events followed by speciation characterized the early biogeographical history of the family**

The results of the AAR analyses and the divergence time estimation showed that the most likely place of origin of the family is the Indo-Pacific region around 112-152 mya in the Cretaceous (Fig. 3). Migration then occurred from there in two directions: North-east and west (Supplemental Data 1). One species of the earliest divergent genus, namely *S. sepioidea,* is nowadays present in Caribbean waters. However, no strong statistical support was obtained for younger nodes that could inform the possible colonization pathway of this species from the Indo-Pacific to the Caribbean region (Fig. 3). The *Doryteuthis/Lolliguncula* clade also comprises several species currently present in the new world coasts. Inferred age and ancestral distribution of this node indicates that this ancestor could arrive to the Atlantic coast via eastward pacific migration from the Indo-Pacific and through the Atlantic long before the Panamá Isthmus closure (Fig. 3).

**Species delimitation of recent divergent populations within cryptic lineages**

We found six species that comprise putative cryptic lineages (Fig. 2). Five of them had already been shown to present population differentiation in other studies (i.e., *D. pleii, D. pealei* (Herke & Foltz, 2002), *L. brevis* (de Luna Sales et al., 2013), *S. lessoniana* and *U. duvauceli* (Bergman, 2013)) while one, *Sepioteuthis sepioidea,* was detected for the first time with our phylogenetic analyses (Fig. 1). However, the bGMYC analysis supported the divergence of phylogeographic groups within these six species only on the *Cytochrome oxidase* I locus. The difference between the *COI* and *Rho* analyses could be explained by the different effective sample sizes (ESS) of mitochondrial versus nuclear genes. ESS of mitochondrial genes are smaller due to the mitochondria maternal inheritance and its haploidy, allowing loci to reach reciprocal monophyly faster (Funk & Omland, 2003; Moore, 1995). Hence, *COI* usually retrieves trees with higher resolution for recently diverged taxa. Although *16S* is also a mitochondrial gene, it has been shown that this gene is less efficient and has lower levels of taxonomic resolution (Funk & Omland, 2003), thus why the best resolution was retrieved with *COI*. However, it has been showed (J. B.

de L. Sales et al., 2017) that the divergence time of the two clades within *D. pleii,* one comprising northwestern and central Caribbean Atlantic and the other southwestern Atlantic specimens, was not recent and instead about 16 million years ago. These results contradict our divergence time estimation, the latter suggesting a split between *D. plei* and *D. pealeii* no older than 5 mya. Thus, the explanation of a recent divergence between phylogeographic groups and inconsistency between molecular markers wouldn't be supported by previous studies but could agree with our divergence time results.

**Divergence time estimates and ENMs support the Amazon-Orinoco Plume as a driver of diversification in *D. pleii* and *L. brevis***

The species *D. pleii*, *D. pealeii* and *L. brevis*. have two genetically isolated populations, each inhabiting the north and south Atlantic coast along the American continent (de Luna Sales and others, 2013). We found that the environmental conditions for the species tested (i.e., *D. pleii* and *L.brevis)* are suitable in both regions which means that northern and southern populations are occupying the same environmental space (except for the southern population of *D. pleii* in the northern Atlantic, Figure 4). However, our niche modelling projections for both species, show this area as environmentally different and unsuitable, mainly in the mouth of the Amazon, which would support the Plume as a barrier. The sea surface temperatures are strongly influenced by rainfall changes in the Amazon River basin, impacting the river discharge and consequently the sea surface salinity in the Amazon plume (Hu et al., 2004; Vizy & Cook, 2010). Low-salinity surface waters are found as far as 2000 km at an average depth of 20-30m from the mouth (Gouveia et al., 2019; Tyaquiçã et al., 2017; Varona et al., 2019). It is one of the greatest discharges of fresh water and suspended sediments in the world, previously presented as a barrier for the dispersal of other organisms such as bacterioplankton, coral and lionfish (de Souza et al., 2017; Hewson et al., 2006; Luiz et al., 2013).

Previous works on biogeography of loliginid squids (Bartol et al., 2002; de Araujo & Gasalla, 2018; Şen, 2005; Zaragoza et al., 2015) have also hypothesized that physicochemical changes in sea water triggered by events such as river drainage may serve as barriers to dispersal for adults, paralarvae and eggs in these animals. Moreover, salinity has been suggested to influence the distribution of squids, due to the narrow salinity ranges that early developmental stages can tolerate (Cinti et al., 2004).

The establishment of the Amazon River system with its current eastwards drainage to the Atlantic Ocean has been dated to occur in the mid-late Miocene, around 10.6 to 9.7 mya (Figueiredo et al., 2010). However, it was not only until the late Pliocene, around 2.5 mya, when it reached the present size and shape (Hoorn et al., 1995). Similar timeline followed the origin and establishment of the Orinoco River during the late Miocene (Cinti et al., 2004; Hoorn et al., 1995). The geological history of the rivers and their Atlantic drainage precedes and approximates our estimated divergence times of populations of all four cryptic lineages along the western Atlantic coast (i.e., *D. pleii, D. pealeii, L. brevis* and *S. sepioidea*), which resulted in splits younger than 5 mya (Figure 3). Coupled with the ecological niche modeling for *D. pleii* and *L. brevis*, these results support the

hypothesis of the Amazonian Orinoco Plume as a barrier which isolated the species in two populations and prevented posterior gene flow, resulting in their consequent divergence.

## Concluding remarks

Our study provides a robust insight into the historical biogeography of the Loliginidae family. Our phylogenetic analysis coupled with the ancestral area reconstruction confirmed that the origin of this family highly likely occurred in the Indo-Pacific region around 112-152 mya, in the Cretaceous. Additionally, we were able to confirm the monophyly of all currently recognized species, to reveal the placement of Lolliguncula panamensis within the Doryteuthis clade and to find what seems to be a possible new cryptic lineage (i.e. *Sepioteuthis sepioidea)*. Further on, our divergence time estimates and ecological niche analyses uncovered the mechanisms and evolutionary history behind the genetic isolation of *D. pleii* and *L. brevi*s populations, with the Amazon-Orinoco Plume playing a very relevant role in the diversification of these species and possibly the other ones distributed along the western Atlantic coast. Our results encourage further studies of the impact that the increase of such barrier may have not only on squids but also in other species.

## DATA AVAILABILITY STATEMENT

All data used in this study was downloaded from open sources as listed in the Methods section and the Supplementary Table 1. The supporting data generated here (i.e. alignments, raw phylogenetic trees, coordinates, layers, occurrences, R script) are available from the Dryad Digital Repository

# REFERENCES

Amante, C., & Eakins, B. W. (2009). *ETOPO1 arc-minute global relief model: Procedures, data sources and analysis*.

Amor, M. D., Norman, M. D., Cameron, H. E., & Strugnell, J. M. (2014). Allopatric speciation within a cryptic species complex of Australasian octopuses. *PLoS One*, *9*(6), e98982.

Anderson, F. E. (2000). Phylogenetic relationships among loliginid squids (Cephalopoda: Myopsida) based on analyses of multiple data sets. *Zoological Journal of the Linnean Society*, *130*(4), 603–633.

Armonies, W. (2001). What an introduced species can tell us about the spatial extension of benthic populations. *Marine Ecology Progress Series*, *209*, 289–294.

Barrett, L. (2017). *Habitat preferences and fitness consequences for fauna associated with novel marine environments*.

Bartol, I., Mann, R., & Vecchione, M. (2002). Distribution of the euryhaline squid Lolliguncula brevis in Chesapeake Bay: Effects of selected abiotic factors. *Marine Ecology Progress Series*, *226*, 235–247.

Bergman, A. M. (2013). Phylogeography of Sepioteuthis lessoniana (the bigfin reef squid) and Uroteuthis duvauceli (the Indian squid). *Honors Theses, Southern Illinois University, Carbondale, United States of America*.

Bierne, N., Bonhomme, F., & David, P. (2003). Habitat preference and the marine-speciation paradox. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1522), 1399–1406.

Bowen, B. W., Rocha, L. A., Toonen, R. J., & Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends in Ecology & Evolution*, *28*(6), 359–366.

Brakoniecki, T. F. (1987). *A generic revision of the family Loliginidae (Cephalopoda: Myopsida) based primarily on the comparative morphology of the hectocotylus (zoogeography)*.

Brandley, M. C., Guiher, T. J., Pyron, R. A., Winne, C. T., & Burbrink, F. T. (2010). Does dispersal across an aquatic geographic barrier obscure phylogeographic structure in the diamond-backed watersnake (Nerodia rhombifer)? *Molecular Phylogenetics and Evolution*, *57*(2), 552–560.

Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., Thuiller, W., Fortin, M., Randin, C., & Zimmermann, N. E. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, *21*(4), 481–497.

Broennimann, O., Petitpierre, B., Randin, C., Engler, R., Breiner, F., D'Amen, M., Pellissier, L., Pottier, J., Pio, D., & Mateo, R. G. (2015). ecospat: Spatial ecology miscellaneous methods. *R Package Version*, *1*.

Cinti, A., Barón, P. J., & Rivas, A. L. (2004). The effects of environmental factors on the embryonic survival of the Patagonian squid Loligo gahi. *Journal of Experimental Marine Biology and Ecology*, *313*(2), 225–240.

Clarke, M., & Maddock, L. (1988). Statoliths of fossil coleoid cephalopods. In *Paleontology and neontology of cephalopods* (pp. 153–168). Elsevier.

Clarke, M. R., & Fitch, J. E. (1979). Statoliths of Cenozoic teuthoid cephalopods from North America. *Palaeontology*, *22*(2), 479–511.

Costello, M. J., Tsai, P., Wong, P. S., Cheung, A. K. L., Basher, Z., & Chaudhary, C. (2017). Marine biogeographic realms and species endemicity. *Nature Communications*, *8*(1), 1–10.

Cowen, R. K., & Sponaugle, S. (2009). *Larval dispersal and marine population connectivity*.

Cowman, P. F., & Bellwood, D. R. (2013). Vicariance across major marine biogeographic barriers: Temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1768), 20131541.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates. http://www.joelvelasco.net/teaching/2890/coyneorr04-speciationch1.pdf

de Araujo, C. C., & Gasalla, M. A. (2018). Distribution patterns of loliginid squid paralarvae in relation to the oceanographic features off the South Brazil Bight (22°–25° S). *Fisheries Oceanography*, *27*(1), 63–75.

de Luna Sales, J. B., Shaw, P. W., Haimovici, M., Markaida, U., Cunha, D. B., Ready, J., Figueiredo-Ready, W. M., Schneider, H., & Sampaio, I. (2013). New molecular phylogeny of the squids of the family Loliginidae with emphasis on the genus Doryteuthis Naef, 1912: Mitochondrial and nuclear sequences indicate the presence of cryptic species in the southern Atlantic Ocean. *Molecular Phylogenetics and Evolution*, *68*(2), 293–299.

de Souza, J. N., Nunes, F. L., Zilberberg, C., Sanchez, J. A., Migotto, A. E., Hoeksema, B. W., Serrano, X. M., Baker, A. C., & Lindner, A. (2017). Contrasting patterns of connectivity among endemic and widespread fire coral species (Millepora spp.) in the tropical Southwestern Atlantic. *Coral Reefs*, *36*(3), 701–716.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*(1), 1–8.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969–1973.

Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, *1*(4), 330–342.

Figueiredo, J., Hoorn, C., Van der Ven, P., & Soares, E. (2010). Late Miocene onset of the Amazon River and the Amazon deep-sea fan: Evidence from the Foz do Amazonas Basin: Reply. *Geology*, *38*(7), e213–e213.

Foggo, A., Attrill, M. J., Frost, M. T., & Rowden, A. A. (2003). Estimating marine species richness: An evaluation of six extrapolative techniques. *Marine Ecology Progress Series*, *248*, 15–26.

Frolova, A., & Miglietta, M. P. (2020). Insights on bloom forming jellyfish (Class: Scyphozoa) in the Gulf of Mexico: Environmental tolerance ranges and limits suggest differences in habitat preference and resistance to climate change among congeners. *Frontiers in Marine Science*, *7*, 93.

Funk, D. J., & Omland, K. E. (2003). Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 397–423.

Gouveia, N., Gherardi, D., & Aragão, L. (2019). The role of the Amazon River plume on the intensification of the hydrological cycle. *Geophysical Research Letters*, *46*(21), 12221–12229.

Hachich, N. F., Bonsall, M. B., Arraut, E. M., Barneche, D. R., Lewinsohn, T. M., & Floeter, S. R. (2015). Island biogeography: Patterns of marine shallow-water organisms in the Atlantic Ocean. *Journal of Biogeography*, *42*(10), 1871–1882.

Heled, J., & Drummond, A. J. (2010). Bayesian Inference of Species Trees from Multilocus Data. *Molecular Biology and Evolution*, *27*(3), 570–580.

Herke, S., & Foltz, D. (2002). Phylogeography of two squid (Loligopealei and L. Plei) in the gulf of mexico and northwestern atlantic ocean. *Marine Biology*, *140*(1), 103–115.

Hewson, I., Capone, D. G., Steele, J. A., & Fuhrman, J. A. (2006). Influence of Amazon and Orinoco offshore surface water plumes on oligotrophic bacterioplankton diversity in the west tropical Atlantic. *Aquatic Microbial Ecology*, *43*(1), 11–22.

Hoorn, C., Guerrero, J., Sarmiento, G. A., & Lorente, M. A. (1995). Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology*, *23*(3), 237–240.

Hu, C., Montgomery, E. T., Schmitt, R. W., & Muller-Karger, F. E. (2004). The dispersal of the Amazon and Orinoco River water in the tropical Atlantic and Caribbean Sea: Observation from space and S-PALACE floats. *Deep Sea Research Part II: Topical Studies in Oceanography*, *51*(10–11), 1151–1171.

Jereb, P., & Roper, C. F. (2010). *Cephalopods of the world. An annotated and illustrated catalogue of cephalopod species known to date. Volume 2. Myopsid and Oegopsid Squids*. FAO.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Kolis, K. M., & Lieberman, B. S. (2019). Using GIS to examine biogeographic and macroevolutionary patterns in some late Paleozoic cephalopods from the North American Midcontinent Sea. *PeerJ*, *7*, e6910.

Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, *29*(6), 1695–1701.

Laughlin, R. A., & Livingston, R. J. (1982). Environmental and trophic determinants of the spatial/temporal distribution of the brief squid (Lolliguncula brevis) in the Apalachicola estuary (North Florida, USA). *Bulletin of Marine Science*, *32*(2), 489–497.

Lindgren, A. R., Pankey, M. S., Hochberg, F. G., & Oakley, T. H. (2012). A multi-gene phylogeny of Cephalopoda supports convergent morphological evolution in association with multiple habitat shifts in the marine environment. *BMC Evolutionary Biology*, *12*(1), 129.

Losos, J. B. (1996). Phylogenetic perspectives on community ecology. *Ecology*, *77*(5), 1344–1354.

Luiz, O. J., Floeter, S. R., Rocha, L. A., & Ferreira, C. E. (2013). Perspectives for the lionfish invasion in the South Atlantic: Are Brazilian reefs protected by the currents? *Marine Ecology Progress Series*, *485*, 1–7.

Maddison, W., & Maddison, D. (2007). Mesquite: A modular system for evolutionary analysis. 2011. *See Http://Mesquiteproject. Org*.

Martins, R., Perez, J., & Schettini, C. (2006). The squid Loligo plei around Santa Catarina Island, southern Brazil: Ecology and interactions with the coastal oceanographic environment. *Journal of Coastal Research*, 1284–1289.

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. 1–8.

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution*, *49*(4), 718–726.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS Biol*, *9*(8), e1001127.

Okutani, T. (1995). *Cuttlefish and squids of the world in color*. National Cooperative Association of Squid Processors.

Patarnello, T., Bargelloni, L., Varotto, V., & Battaglia, B. (1996). Krill evolution and the Antarctic ocean currents: Evidence of vicariant speciation as inferred by molecular data. *Marine Biology*, *126*(4), 603–608.

Phillips, S., Dudík, M., & Schapire, R. (2010). *Maxent software for species habitat modeling: Version 3.3. 3*.

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175.

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. *Systematic Biology*, *55*(4), 595–609.

Pyron, R. A., & Burbrink, F. T. (2010). Hard and soft allopatry: Physically and ecologically mediated modes of geographic speciation. *Journal of Biogeography*, *37*(10), 2005–2015.

Rambaut, A, & Drummond, A. (2007). TreeAnnotator. *Program and Documentation Distributed by the Author. Website Http://Beast. Bio. Ed. Ac. Uk/TreeAnnotator*.

Rambaut, A, & Drummond, A. (2015). *LogCombiner v1. 8.2*.

Rambaut, Andrew, Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901.

Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, *57*(1), 4–14.

Reid, N. M., & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: A Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology*, *12*(1), 196.

Rocha, L. A., Robertson, D. R., Roman, J., & Bowen, B. W. (2005). Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1563), 573–579.

RStudio Team. (2020). RStudio: Integrated development for R. RStudio Inc.(2015). *Acknowledgements This Work Was Supported by Fundação Para a Ciência e Tecnologia (FCT), Azores*.

Sales, J. B. de L., Rodrigues-Filho, L. F. da S., Ferreira, Y. do S., Carneiro, J., Asp, N. E., Shaw, P. W., Haimovici, M., Markaida, U., Ready, J., & Schneider, H. (2017). Divergence of cryptic species of Doryteuthis plei Blainville, 1823 (Loliginidae, Cephalopoda) in the Western Atlantic Ocean is associated with the formation of the Caribbean Sea. *Molecular Phylogenetics and Evolution*, *106*, 44–54.

Sales, J. B. L., Markaida, U., Shaw, P. W., Haimovici, M., Ready, J. S., Figueredo-Ready, W. M., Angioletti, F., Carneiro, M. A., Schneider, H., & Sampaio, I. (2014). Molecular phylogeny of the Genus Lolliguncula Steenstrup, 1881 based on nuclear and mitochondrial DNA sequences indicates genetic isolation of populations from North and South Atlantic, and the possible presence of further cryptic species. *PloS One*, *9*(2), e88693.

Schoener, T. W. (1970). Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, *51*(3), 408–418.

Seibel, B. A. (2007). On the depth and scale of metabolic rate variation: Scaling of oxygen consumption rates and enzymatic activity in the Class Cephalopoda (Mollusca). *Journal of Experimental Biology*, *210*(1), 1–11.

Şen, H. (2005). Incubation of European Squid (Loligo vulgaris Lamarck, 1798) eggs at different salinities. *Aquaculture Research*, *36*(9), 876–881.

Simkanin, C., Carlton, J. T., Steves, B., Fofonoff, P., Nelson, J. C., Clarke Murray, C., & Ruiz, G. M. (2019). Exploring potential establishment of marine rafting species after transoceanic long-distance dispersal. *Global Ecology and Biogeography*, *28*(5), 588–600.

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. (2014). The drivers of tropical speciation. *Nature*, *515*(7527), 406–409.

Sobrino, I., Silva, L., Bellido, J., & Ramos, F. (2002). Rainfall, river discharges and sea temperature as factors affecting abundance of two coastal benthic cephalopod species in the Gulf of Cadiz (SW Spain). *Bulletin of Marine Science*, *71*(2), 851–865.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., & Lourie, S. A. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, *57*(7), 573–583.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Strugnell, J., Jackson, J., Drummond, A. J., & Cooper, A. (2006). Divergence time estimates for major cephalopod groups: Evidence from multiple genes. *Cladistics*, *22*(1), 89–96.

Strugnell, J., & Nishiguchi, M. K. (2007). Molecular phylogeny of coleoid cephalopods (Mollusca: Cephalopoda) inferred from three mitochondrial and six nuclear loci: A comparison of alignment, implied alignment and analysis methods. *Journal of Molluscan Studies*, *73*(4), 399–410.

Tanner, R. A. (2018). *A Phylogennomic Enquiry Into Metazoan Macroevolutionary Dynamics*.

Tournayre, O., Leuchtmann, M., Filippi-Codaccioni, O., Trillat, M., Piry, S., Pontier, D., Charbonnel, N., & Galan, M. (2019). In silico and empirical evaluation of twelve COI & 16S metabarcoding primer sets for insectivorous diet analyses. *BioRxiv*, 742874.

Tyaquiçã, P., Veleda, D., Lefèvre, N., Araujo, M., Noriega, C., Caniaux, G., Servain, J., & Silva, T. (2017). Amazon plume salinity response to ocean teleconnections. *Frontiers in Marine Science*, *4*, 250.

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*, *21*(2), 272–281.

Ulloa, P. M., Hernandez, C. E., Rivera, R. J., & Ibanez, C. M. (2017). Historical biogeography of the squids from the family Loliginidae (Teuthoidea: Myopsida). *Latin American Journal of Aquatic Research*, *45*(1), 113–129.

Varona, H., Veleda, D., Silva, M., Cintra, M., & Araujo, M. (2019). Amazon River plume influence on Western Tropical Atlantic dynamic variability. *Dynamics of Atmospheres and Oceans*, *85*, 1–15.

Vecchione, M. (1991). Observations on the paralarval ecology of a euryhaline squid, Lolliguncula brevis (Cephalopoda; Loliginidae). *Fishery Bulletin*.

Vizy, E. K., & Cook, K. H. (2010). Influence of the Amazon/Orinoco Plume on the summertime Atlantic climate. *Journal of Geophysical Research: Atmospheres*, *115*(D21).

Wani, R. (2011). Sympatric speciation drove the macroevolution of fossil cephalopods. *Geology*, *39*(11), 1079–1082.

Warren, W. C., Kuderna, L., Alexander, A., Catchen, J., Pérez-Silva, J. G., López-Otín, C., Quesada, V., Minx, P., Tomlinson, C., Montague, M. J., Farias, F. H. G., Walter, R. B., Marques-Bonet, T., Glenn, T., Kieran, T. J., Wise, S. S., Wise, J. P., Waterhouse, R. M.,

& Wise, J. P. (2017). The Novel Evolution of the Sperm Whale Genome. *Genome Biology and Evolution*, *9*(12), 3260–3264. https://doi.org/10.1093/gbe/evx187

Wiens, J. J. (2004). Speciation and ecology revisited: Phylogenetic niche conservatism and the origin of species. *Evolution*, *58*(1), 193–197.

Wu, R., Zhang, H., Liu, J., Niu, S., Xiao, Y., & Chen, Y. (2018). DNA barcoding of the family Sparidae along the coast of China and revelation of potential cryptic diversity in the Indo-West Pacific oceans based on COI and 16S rRNA genes. *Journal of Oceanology and Limnology*, *36*(5), 1753–1770.

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

Zaragoza, N., Quetglas, A., Hidalgo, M., Alvarez-Berastegui, D., Balbín, R., & Alemany, F. (2015). Effects of contrasting oceanographic conditions on the spatiotemporal distribution of Mediterranean cephalopod paralarvae. *Hydrobiologia*, *749*(1), 1–14.

Zielinski, S., Lee, P., & Pörtner, H.-O. (2000). Metabolic performance of the squid Lolliguncula brevis (Cephalopoda) during hypoxia: An analysis of the critical PO2. *Journal of Experimental Marine Biology and Ecology*, *243*(2), 241–259.

**BIOSKETCH**

Gabrielle Genty has a masters in science from the University of St. Andrews, Scotland, and is currently working as a research field assistant for Flinders University in Australia. Her main interests are in historical biogeography and evolutionary genetics of marine organisms.

**Author contribution:** C.J.P.H., E.A.R. and G.G. conceived the project; and all authors analyzed the data, worked on the figures and wrote the manuscript.