

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS**

Samuel Chagas Bernardes

**The evolutionary history of *Caridina typus*:
phylogeography, species delimitation and an analysis
of the complexity of the Indo-Pacific biogeography**

Belo Horizonte – MG

2018

Samuel Chagas Bernardes

**The evolutionary history of *Caridina typus*:
phylogeography, species delimitation and an analysis of
the complexity of the Indo-Pacific biogeography**

**Dissertação apresentada ao Programa de Pós-
graduação em Zoologia do Instituto de
Ciências Biológicas da Universidade Federal
de Minas Gerais como requisito parcial para
obtenção do grau de Mestre em Zoologia –
Área de concentração em**

Orientador: Almir Rogério Pepato

Belo Horizonte – MG

2018

I

ACKNOWLEDGMENTS

My most sincere thanks:

- to my advisor, Almir Pepato, whose support, both personal and professional, was priceless through this project;
- to Mark de Bruyn, who has conceded me autonomy to conduct this project when I was still an undergraduate and an exchange student in the Bangor University and has indeed become a major source of support and inspiration;
- to my fiancée, Maíra, who has been my main source of support and motivation ever since she entered my life;
- to my brother, Raphael, who was regularly the only thing that took me away from the computer to take a breath through all the work;
- to my intern in Bangor, Sarah Karpati, who has no idea of how important she was to my progress;
- to the mentors I had through my academic journey who contributed at some point to my formation: Adalberto, Almir, André, Bernadete, Fabrício, Fernando, Flávio, Gary, Gustavo, Kat, Maria Amélia, Mário, Mark, Teofânia and Ubirajara;
- to the masters I've never met but who represent the shoulders of giants on which I stood to see more and farther: Darwin, Wallace, Huxley, Avise, Dobzhansky, Fisher, Gould, Nei, Kimura, Tajima, Felsenstein, Hennig, Haeckel, Haldane, Maynard Smith, Kingman, Platnick, Nelson among many others;
- to the CNPq for the investment, both during the travel that originated the project (Science Without Borders) and during its execution in Brazil;
- to my friends: Adam, Camilla, Carol, Ênio, Larissa, Leonardo, Marco, Michael, Pedro, Raphael, Yasmim and Zoe, and why not include Almir and Mark, who have become great friends, and Maíra, who is my friend for life.

“We are like dwarves perched on the shoulders of giants, and thus we are able to see more and farther than the latter.”

- Isaac Newton

“(…) The ascent of man is not made by lovable people. It is made by people who have two qualities: an immense integrity, and at least a little genius.”

- Jacob Bronowski

RESUMO

O Oceano Índico inclui em sua área de influência nove *hotspots* de biodiversidade, incluindo alguns dos mais icônicos (e, portanto, mais estudados) centros de endemismo do mundo. Historicamente, a região tem servido como um prolífico laboratório de estudos geológicos e biogeográficos e, avaliando a distribuição da diversidade, nos foi possível inferir sobre a história biológica da região e sobre os processos que a influenciam, especialmente, no que diz respeito ao endemismo. O estudo de espécies não-marinhas com ampla distribuição através do Oceano Índico como *Caridina typus* H. Milne-Edwards, 1837 pode esclarecer muitos desses dilemas acerca da história recente dessa área e da dinâmica evolutiva de espécies vageis. O objetivo primário desse estudo é, portanto, a avaliação da história evolutiva das populações de *C. typus* bem como a apuração de sua filogeografia e status taxonômico. Foram amplificados cinco marcadores (dois mitocondriais e três nucleares) de 129 indivíduos ao longo da região Indo-Pacífica (área de ocorrência da espécie). No capítulo 3 está a filogenia datada e a inferência de uma história evolutiva para o complexo de espécies. Ainda nesse capítulo, é sugerida a delimitação de três espécies dentro de *C. typus* feita através de três métodos diferentes. Ao mostrar, pela primeira vez na literatura, um resultado consistente entre vários métodos diferentes, faz-se também uma breve crítica sobre a delimitação de espécies baseada em marcadores mitocondriais. Com base no cenário teórico em que essa pesquisa está inserida, um segundo objetivo foi estabelecido: uma revisão e meta-análise de estudos biogeográficos moleculares realizados na região do Oceano Índico. Esse oceano tem uma complexa história geológica relacionada com a fragmentação de supercontinentes, paleo-oceanos e com a travessia da Índia através dos hemisférios. No entanto, o estudo biogeográfico da região foi mais influenciado pelas escolas de pensamento do que pela evidência real da biodiversidade. Com o uso de novas metodologias associadas ao desenvolvimento da biologia molecular, pode-se elucidar muitas das discussões a respeito das áreas de endemismo do Oceano Índico. No capítulo 2 (organizado assim por motivos práticos), foi realizada uma extensiva revisão da literatura acerca da história geológica e biológica da região e 7 trabalhos filogenéticos mais ou menos geograficamente inclusivos foram reanalisados. Os resultados mostram que a história biogeográfica da região é complexa e envolve diversos eventos de vicariância, dispersão (curta e transoceânica) e que a área ainda tem muito a oferecer em termos de conhecimentos sobre processos evolutivos.

ABSTRACT

The Indian Ocean comprises nine hotspots of biodiversity in its area of influence, including some of the world's most iconic (and thus most studied) centres of endemism. Historically, the area serves as a prolific laboratory for geologic and biogeographic studies and by assessing the biodiversity distribution it became possible to infer about the biological history of the region and the processes that influence it, particularly in what refers to endemism. The study of non-marine species widely distributed across the Indian Ocean such as *Caridina typus* H. Milne-Edwards, 1837 may shed some light on many of these dilemmas on the recent history of this area and of the evolutionary dynamics of vagile species. The primary objective of this study is, therefore, the investigation of the evolutionary history of the *C. typus*' populations and to canvass their relations and taxonomic status. Five molecular markers were amplified (two mitochondrial and three nuclear) for 129 individuals across the Indo-Pacific region (species' area of occurrence). *Caridina typus*' dated phylogeny can be found in chapter 3 with an inference of the evolutionary history of the species complex. Still in chapter 3, three species were suggested by delimitation methods in *C. typus* by three different methods. By showing, for the first time in literature, an agreement between various methods, a brief critique on the species delimitation based on mitochondrial markers was also made. Based on the theoretical scenario in which this research is inserted, a second aim was established: a review and meta-analysis of the biogeographical molecular studies done across the Indian Ocean. That ocean has a complex geological history, intimately related to the fragmentation of supercontinents, paleo-oceans and the India's drift through the hemispheres. However, biogeographic studies there have been more influenced by biogeographic schools than by real evidence from biodiversity. With the new methods associated to the development of the molecular biology one can elucidate a great part of the discussions on Indian Ocean's areas of endemism. In chapter 2 (organised this way due to practical reasons), the literature about the geological and biological history of the region was extensively reviewed and 7 phylogenetic works more or less geographically inclusive were re-analysed. The results show that the biogeographic history of the region is complex and involves various events of vicariance, dispersal (both short and long-distance) and that the area still has much to offer in terms of knowledge on evolutionary processes.

FIGURE LIST

Figure 1.1: Lateral view of an atyid highlighting the family's particular features.

Figure 1.2: Distribution of the atyids and of the genus *Caridina* – and thus of *Caridina typus*.

Figure 2.1: Geological reconstruction of the Indian Ocean continents through the Mesozoic and Cenozoic eras.

Figure 2.2: Ancestral area reconstruction and biogeographic history.

Figure 3.1: Sampling map of the Indo-Australian Archipelago (IAA) and Eastern Asia (a), with detail for Philippine islands (b); Western Indian Ocean (c) and South Pacific (d).

Figure 3.2: Time-calibrated phylogenetic and phylogeographic relationships in *Caridina typus* rooted using *C. opaensis*.

Figure 3.3: Mitochondrial 16S haplotype network.

Figure 3.4: Extended Bayesian skyline plot for mtDNA and multiloci data for each main clade of *Caridina typus*.

Figure 3.5: An illustrated suggestion for the biogeographic history of *Caridina typus*.

Supplementary Figure S1: 28S alignment through secondary structure for the 18 haplotypes represented in the network.

Supplementary Figure S2: Multispecies mitochondrial 16S Bayesian *Caridina* tree.

Supplementary Figure S3: Haplotype networks for all five markers.

Supplementary Figure S4: STACEY matrix.

Supplementary Figure S5: RASP's BayArea full result.

Supplementary Figure S6: BEAST discrete phylogeography result with collapsed branches per range.

Supplementary Figure S7: Partial COI (167 nucleotide) maximum likelihood and neighbour-joining tree including Fujita et al.'s (2016) sequences.

ABBREVIATION LIST

°C: Celsius degrees

ANM: Anonymous Nuclear Marker

bp: base pair

COI: Subunity I of the cytochrome c oxidase

DNA: Deoxyribonucleic acid

IAA: Indo-Australian Archipelago

K-Pg: Cretaceous–Paleogene

Ma: Mega annum (anna), i.e., Million(s) years

min: Minute(s)

MIOJet: Miocene Indian Ocean Equatorial Jet

mm: Millimetre(s)

OTU: Operational taxonomic unity(ies)

s: Second(s)

CONTENTS

TITLE	I
ACKNOWLEDGMENTS	II
RESUMO.....	VI
ABSTRACT.....	VII
FIGURE LIST.....	VIII
ABBREVIATION LIST	IX
CONTENTS.....	X
Chapter 1: General introduction.....	1
1.1 Taxonomy and evolution of the Atyidae (Decapoda: Caridea).....	1
1.2 Molecular systematics and taxonomy	3
1.3 Phylogeography and population genetics.....	4
1.4 Thesis justification and aims	5
Chapter 2: A brief review and analysis of the Indian Ocean biogeography – A highway for continental species	10
2.1 Abstract	10
2.2 Introduction	10
2.3 Overview of the Indian Ocean’s geology.....	12
2.3.1 On the Eastern Indian Ocean	15
2.3.2 On the Western Indian Ocean	15
2.4 The Indian Ocean’s biogeography	16
2.5 Materials and methods	20
2.5.1 Studies chosen.....	20
2.5.2 Phylogenetic analyses	21
2.5.3 Biogeographic analyses	24
2.6 Results	24
2.6.1 Ancient taxa diversification	24
2.6.2 Volant taxa diversification.....	27
2.6.4 Recent taxa diversification	28
2.7 Discussion	28
2.7.1 On an isolated India and its relationship with Seychelles and Madagascar	28
2.7.2 On the colonisation of the Indian Ocean islands	29
2.8 Conclusion.....	31
Chapter 3: The complex evolutionary history and phylogeography of <i>Caridina typus</i> (Crustacea: Decapoda): long-distance dispersal and cryptic allopatric speciation with comments on species delimitation	32
3.1 Abstract	32

3.2	Introduction	33
3.2	Materials and methods	34
3.2.1	Sampling	34
3.2.2	DNA – extraction, amplification and sequencing.....	36
3.2.3	Phylogenetic analyses	36
3.2.4	Analyses with Datasets: priors and model setting	37
3.2.5	Species delimitation analyses	39
3.2.6	Demographic and phylogeographic analyses.....	39
3.3	Results	40
3.3.1	DNA – extraction, amplification and sequencing.....	40
3.3.2	Phylogenetic analyses	40
3.3.3	Species delimitation analyses	41
3.3.4	Demographic and phylogeographic analyses.....	44
3.4	Discussion	47
3.5	Conclusion.....	53
	Bibliography	54
	Appendix.....	77

Chapter 1: General introduction

This is a work of phylogeography with comments on taxonomy. Phylogenetics and population genetics are intrinsic to phylogeography as establishing the relations between and inside populations is paramount to demonstrate a taxon's macro and microevolutionary aspects. Therefore, just as researches on phylogeography and population genetics frequently lead to questioning and/or altering the current biological classification, the observations of the natural populations made by taxonomists may suggest phylogeographic studies to assess their connections.

Even though this study is about the use of genetic tools to infer a taxon's distribution history, molecular systematics is needed to study the past and trends in its current populations. That makes it impossible to overlook a discussion on its taxonomic status. Also, this work brings preliminary results of a review on biogeographic works accomplished through the Indian Ocean area. Such review is important to evidence the main rules and processes involved with the distribution patterns we find today.

1.1 Taxonomy and evolution of the Atyidae (Decapoda: Caridea)

The Atyidae are a family of freshwater shrimps found across all continents, but Antarctica (Williams 1980). Their most obvious synapomorphy are the setae found on the tip of the chelae of their maxillipeds; those are thin, hair-like, and are used to scrape food off the substrate (Bruce 1992). Furthermore, the atyids can be characterised by the presence of a supraorbital spine and exopods on the pereopods (which are uniramous among most Decapoda; see fig. 1.1) (Williams 1980). Both molecular (Porter et al. 2005) and morphologic (Felgenhauer & Abele 1985) data point to an early divergence of the family inside the Caridea. Porter et al. (2005) suggest that the divergence between atyids and other Caridea happened at some point in the Late Permian and the diversification within the family happened in the Early Jurassic.

Atyids are usually smaller than 35 mm (with a few exceptions; see Chace 1983) and are found in all types of freshwater bodies in addition to anchialine formations and caves, plus some species show an estuarine reproduction (Davie 2002). There are no extant marine atyids (Huxley 1880; Fryer 1977), but their life histories vary a lot according to the species or genus. Such diversity seems to be intrinsic to the morphology of the egg (Jalihal et al. 1994; Hancock

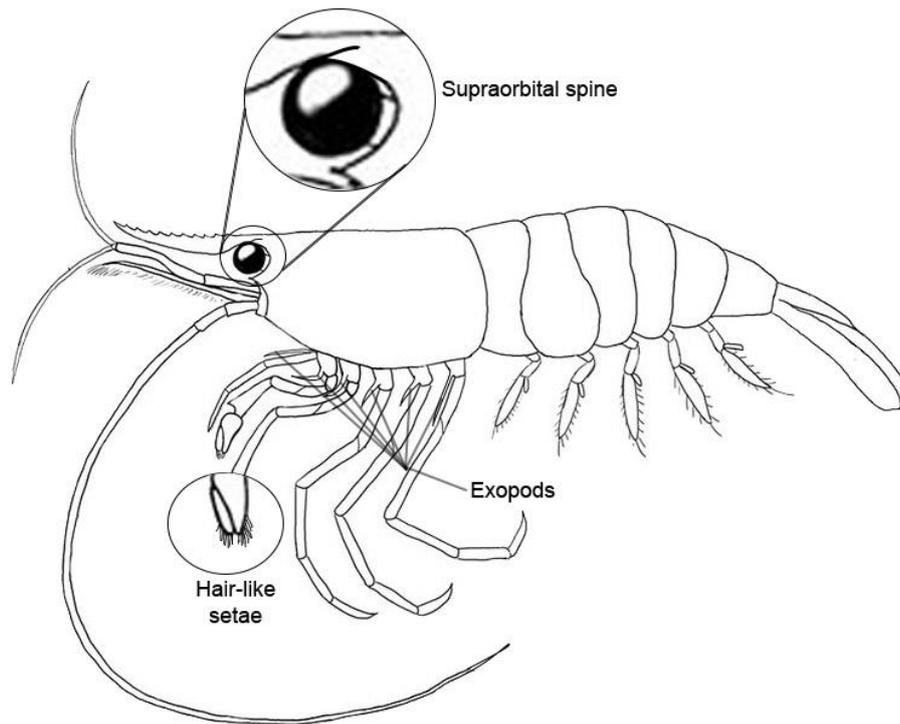


Figure 1.1: Lateral view of an atyid highlighting the family's particular features (adapted from Williams 1980)

1998): small eggs indicate planktonic larvae with high tolerance to salinity (often belonging to species with estuarine breeding), large eggs belong to species with direct development and low tolerance to salinity and intermediate-sized eggs point out to species with transitional features in both development and salinity tolerance (Hayashi & Hamano 1984; Hancock et al. 1998; Shy et al. 2001; Fièvet & Eppe 2002; Short & Doumenq 2003).

Given their reduced size and abundance in freshwater environments, atyids have major ecosystemic roles. First, as Yam and Dudgeon (2006) showed, species in the genus *Caridina* H. Milne-Edwards, 1837 show a higher production rate if compared to similar crustaceans. In addition, in the family are some of the most important algae (March & Pringle 2003) and particulate matter consumers, whose absence may unbalance a system's biomass and sedimentation levels (Yam & Dudgeon 2005; de Souza & Moulton 2005). Besides, the atyids are an important prey in their ecosystems (Covich et al. 1999).

The biogeography of the Atyidae clearly involves both vicariance (given the family's age, older than most of the continental separation events) and dispersal (thanks to the diversity in salinity tolerance and history of life) (Bănărescu 1990; von Rintelen et al. 2007, 2012). Also, their classification is historically troublesome (Huxley 1880; Fryer 1977) and has been going

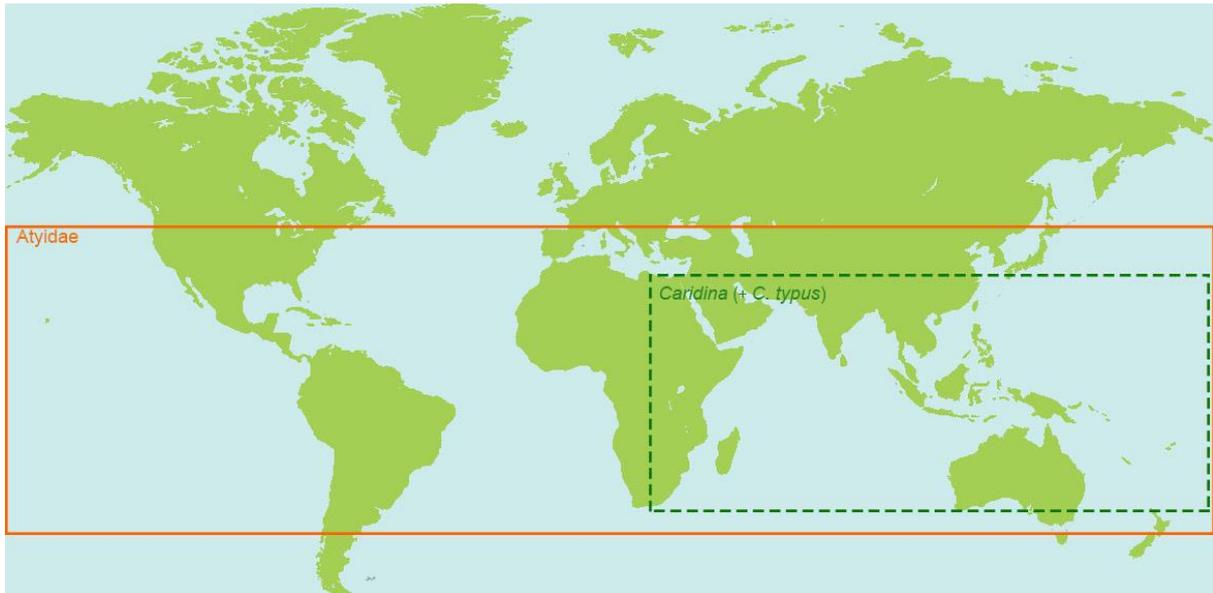


Figure 1.2: Distribution of the atyids and of the genus *Caridina* – and thus of *Caridina typus*.

through many changes recently (e.g. Page et al. 2005; Porter et al. 2005; von Rintelen et al. 2012; Jurado-Rivera et al. 2016). *Caridina* is probably the most problematic genus: taxonomically it has been largely neglected, with over 300 described species and dozens of ambiguities (Fransen 2015); biogeographically, with species distributed across the paleotropical and part of the palearctic regions (see fig. 1.2), it has been shown to have a very complex and active history with multiple events of dispersal, vicariance and environmental invasion (von Rintelen et al. 2012; Jurado-Rivera et al. 2017). The phylogeography and taxonomic assessment of the Atyidae are thus particularly interesting to the evolutionary dynamics of freshwater species.

1.2 Molecular systematics and taxonomy

Molecular systematics is the field of systematics that uses molecular data, mainly DNA sequences, to clarify the evolutionary relations between the organisms and evaluate the current taxonomy (Hillis et al. 1996). This branch arose in the 1960's when the development of techniques in molecular biology brought grandiose promises of resolution of the 'tree of life' (Zuckerandl & Pauling 1965). Naturally, such promises found equally grandiose obstacles and propelled methodological advances. It also inflamed philosophical discussions, which were revived in the last decade with the emergence of the Barcode of Life project, an attempt to reduce species identification to a DNA analysis (Hebert et al. 2003).

One of the main concerns of the Barcode project is the identification of cryptical species – i. e., two or more species that have been historically described under a single scientific name due to morphological indistinguishability (Bickford et al. 2006) –, which is problematic with a morphological approach (Hebert et al. 2003). Thereby, the Barcode brought new promises and expectations on the use of the molecular data (Blaxter 2003; Proudlove & Wood 2003; reviewed by Bickford et al. 2006), but the debate on how to assess the taxonomic status of new genetically determined ‘species’ is still a hot topic due to favouritism of one or other method (Dunn 2003; Hebert et al. 2003; Lipscomb et al. 2003; Scotland et al. 2003; Tautz et al. 2003; Blaxter 2004). To assess each morphologic and molecular method’s pros and cons generated debates in both systematics and conservation (ex. Daugherty et al. 1990; Geller 1999; Hay et al. 2010) as well as in the way that we understand old and new types of data (ex. Will & Rubinoff 2004; Collins & Cruickshank 2014).

This scenario created a demand for methods that allowed the evaluation of molecular phylogenies in order to delimit ‘species’ or operational taxonomic unities (OTU) especially for modest datasets based on the Barcode approach (Carstens et al. 2013). Some of these methods, essentially linked to the Barcode, require no more than one locus (ex. Fujisawa & Barraclough 2013). There are a considerable number of methods nowadays and they may be easily compared as they deal with different evolutionary problems, properties and concepts (Carstens et al. 2013; Satler et al. 2013; see de Queiroz 2007). Carstens et al. (2013) evidenced that many studies present incongruences between different methods and justify them by showing how the methods’ premises were violated. Consequently, the authors propose that to make use of several methods and assess their concordance is a good way to deal with delimitation.

1.3 Phylogeography and population genetics

Ortmann (1994) asserted that to establish the relations and affinities between taxa is primary to understand and find out the causes for the geographical distribution. The author is referring to the identification and classification of species that have a wide distribution and are often recognised as cryptical species (Bohonak & Jenkins 2003) or as result of homoplasy (Bossuyt et al. 2004). Biogeographical research and hypotheses are inherent to the phylogeny as the reconstruction of the relationships between populations raises and addresses questions about the framework related to them (Emerson 2002). It means that scenarios where the taxonomy is not directly connected to the phylogeny can (and frequently will) lead biogeographical studies to major mistakes or, at least, to dead ends in the understanding.

While phylogeny has been associated to historical biogeography, population genetics is more used in ecological biogeography. The emergence of population genetics represents an expressive advance in evolutionary biology, being the core of the modern synthesis (Beatty 1986) and presenting mathematical-scientific rigour for the area (Provine 1988). Dobzhansky's *Genetics and the Origin of Species* (1937) disturbed the population genetics by contradicting the premise that populations are genetically homogeneous, which was assumed by most geneticists at the time; Dobzhansky demonstrated that populations present not only a wide diversity but also notable distinctions between subpopulations. The analysis of those different gene pools in order to describe the relations between alleles/haplotypes across the geographical distribution led to the development of a whole new area, capable to deal with both historical and ecological biogeography: phylogeography (Avice et al. 1987; Avice 2000).

Phylogeography is an important field of evolutionary biology because it is able to connect historical and ecological biogeography as well as phylogenetics and population genetics (Avice 2004). Thanks to that ability, phylogeography has influenced the development of biogeographical methods, especially those related to the reconstruction of ancestral distribution ranges (e.g. Ree & Smith 2008; Yu et al. 2010; Landis et al. 2013). It has also allowed a reassessment of the vicariance-dispersal dichotomy that has been kept for decades (Zink et al. 2000). A convenient example is oceanic dispersal: widely neglected after the discovery of plate tectonics, it was re-established through molecular dating and modern biogeographical tools (reviewed by Queiroz 2005).

The palaeotropical region, other than being a scenario for Wallace's first published biogeographical works, presents a rich and troubled geological history. Researches across this region are consequently not rare – particularly those involving the Indo-Australian Archipelago (IAA) – and they still often succeed in offering novel contributions to evolutionary biology. The region was a representative part in recent studies that re-evaluated oceanic dispersal's role in the organismal distribution (Zink et al. 2000; de Queiroz 2005).

1.4 Thesis justification and aims

One of the main reasons for which these taxa are interesting for biogeography and evolutionary biology are the boundaries imposed to their dispersal by both terrestrial and oceanic environments. The interest for freshwater species began early in biogeography: Wallace (1881), Darwin (1859) and Huxley (1880) have commented or discussed the

distribution of the freshwater species in their publications (particularly, ‘Darwin’s Bulldog’ exhibits great interest in a large-scale biogeography of the decapods (Huxley 1880)). John Avise, underlines the relevance of such organisms in his works: freshwater organisms are offered too many types of obstacles and, therefore, they develop the largest amount of intraspecific genetic diversity and present strong evidence for refugia (Avise et al. 1987; Avise 2000, 2004, 2009). Unaware of plate tectonics, Darwin (1859) noticed that freshwater organisms are good models for researching and/or testing vicariant events; besides, he discusses the dispersal possibilities that some species could have developed.

The type species of the genus *Caridina*, *Caridina typus* H. Milne-Edwards, 1837, has already caught scientists’ attention with its wide distribution: contrary to most *Caridina* species, which are restricted to islands or specific landlocked locations, *C. typus*’ range is almost completely congruent with the distribution of the entire genus (Bouvier 1925; Johnson 1960, 1961). One of the reasons for that is perhaps *C. typus*’ intimate connection to the coast, with planktonic larvae and probably estuarine reproduction (Johnson 1963; Suzuki et al. 1993). In spite of these intriguing features, this species has never gone through any populational or phylogeographic studies, though there are suggestions that it is actually a species complex (de Grave 2013).

This study has a two-fold aim: first, I shall use mitochondrial and nuclear loci to infer the evolutionary history of the widely distributed populations of *Caridina typus* while evaluating their taxonomic status. With the relationship between populations clarified, it will be possible to make inferences about their ancestral distribution and how they have achieved their current one. In order to assess *C. typus*’ taxonomic status under a molecular biology approach, I shall include specimens of *Caridina villadolidi* Blanco, 1939 in the analyses. This species has been previously described as a subspecies of *C. typus* and there’s still some difficulty to morphologically distinguish it from *C. typus* (Cai & Ng 2001; Thomas von Rintelen pers. comm.).

Such phylogeographic studies, focused on a single species, have been said to be ‘narrative’ and ‘non-empirical’ in contrast to cladistic biogeography, mainly concerned with describing the history of the areas instead of the taxa (Humphries & Parenti 2001; reviewed by Knapp 2005). However, as Knapp (2005) has put, patterns inferred by such works are valuable even to cladistic biogeography itself: when multiple taxa show congruent patterns, they lead to deeper more enlightening research – and one might add that, if they are too incongruent, they

may also inspire new investigations. A species as *C. typus*, with such large distribution, across different climates and environments, is thus priceless to biogeography. Besides, even though *C. typus* is not a threatened species, it is present in 14 of the 34 current recognised biodiversity hotspots (De Grave 2013; Fransen 2015). Therefore, apart from offering an opportunity of addressing problems related to biodiversity (such species delimitation and understanding of the radiation processes), *C. typus* may also be a candidate for bioindication of both lotic and lentic – and, perhaps, even estuarine – environments.

Secondly, I shall make a brief review on the biogeography of the Indian Ocean concerning terrestrial and freshwater taxa. Initially, the geological history of the cited ocean will be appraised, drawing attention to processes and events that may influence the distribution of non-maritime species. The taxon-range choice was made because patterns and processes involved with the distribution of oceanic species are very different from those regarding terrestrial and freshwater species and are thus difficult to analyse at the same time (see Costello et al. 2017). Then, I will make a meta-analysis across published phylogeographic studies on the taxa of the continents and island bathed by the Indian Ocean: molecular biology works will be remade (corrected and/or complemented where necessary) and submitted to ancestral range inference.

To assist the comprehension of the results and discussions reported here, this dissertation was arranged to have the review and meta-analysis presented before the phylogeographic work on *C. typus*.

On the meta-analysis, the following question was raised:

1. *Are there repetitive large-scale patterns among the distribution of different non-maritime species?*

Can the distribution of different terrestrial/freshwater taxa comprise the same or similar processes? Can predictions be made on how the species are distributed across the Indian Ocean? Do distinct events carry the same importance for different taxa?

Based on this question, the following aims and objectives were defined:

1. *Aim:*

This meta-analysis' aim is to re-analyse published data from animal and plant taxa from across the Indian Ocean and establish a time-calibrated tree and the ancestral distributions in it in order to overlay the histories of both taxa and ocean and assess events and processes that may be involved with the biogeography of terrestrial and freshwater species.

2. *Objectives:*

- a. To obtain time-calibrated trees as precise as possible given the data and information in the literature for each taxon;
- b. To estimate the ancestral distributions based on those trees and on the current distribution of the analysed taxa;
- c. To evaluate which biogeographical hypotheses fit the dates and ancestral ranges contained in the obtained trees.

On the *C. typus* work, several questions were raised:

1. *Is there geographic structure in that species?*

Is there differentiation between *C. typus*' populations? If so, does it reflect their geographic distribution?

2. *How and when did C. typus reach its current distribution?*

Which types of processes and events are involved with this species' distribution? What is the relative importance of vicariance and dispersal for it?

3. *What is the taxonomic status of C. typus?*

Does *Caridina typus* comprise a single species without significant differences between populations or is it in fact a complex of two or more cryptical species?

Based on these questions, the following aims and objectives were defined:

1. *Aim:*

This research's aim is to use both mitochondrial and nuclear data to describe the distribution of the genetic diversity of the *Caridina typus*' populations, reveal whatever patterns of geographic structure, and assess its taxonomic status through molecular systematics methods.

2. *Objectives:*

- a. To test the hypothesis that different populations represent different lineages according to their geographic arrange;
- b. To evaluate the genetic diversity of different regions to assess population structure;
- c. To evaluate the evolutionary history inside and between populations, to test the hypothesis of ancient radiation and deep coalescence and to assess the existence of multiple species inside a complex.

Chapter 2: A brief review and analysis of the Indian Ocean biogeography – A highway for continental species

Authors: Samuel C. Bernardes¹, Almir R. Pepato¹, Fabrício Lapolli², Thomas von Rintelen³, Kristina von Rintelen³, Timothy J. Page^{4,5}, Björn Stelbrink⁶, Mark de Bruyn^{7,8}

¹Laboratório de Acarologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte 31270-901, Brazil. ²Laboratório de Dinâmica Oceânica, Instituto Oceanográfico da Universidade de São Paulo, Praça do Oceanográfico 191, São Paulo 05508-120, Brazil. ³Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, Berlin 10115, Germany. ⁴Australian Rivers Institute, Griffith University, Nathan, Queensland 4111, Australia. ⁵Water Planning Ecology, Queensland Dept. of Science, Information Technology and Innovation, Dutton Park, Queensland 4102, Australia. ⁶Department of Animal Ecology & Systematics, Justus Liebig University, Heinrich-Buff-Ring 26-32 IFZ, Giessen D-35392, Germany. ⁷Macleay Building (A12), Faculty of Science, University of Sydney, Science Road, Sydney NSW 2006, Australia. ⁸School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW 2052, Australia.

2.1 Abstract

The Indian Ocean has been a recurrent target to biogeography due to its unique geology and controversial biological feature: such as areas of high endemism, complex history and mixed elements. Much discussion has been placed on the topic and each of Indian Ocean's landmasses may represent a particular field of study. Here we re-analyse published data on Indian Ocean taxa by constructing time-calibrated phylogenies and inferring ancestral distribution ranges. Our results suggest that India has a mixed history and that it is naïve to expect signs of continued isolation since its break-up from Gondwana. Also, they show that the islands at the African eastern coast are intimately related and that Eastern and Western Indian Ocean have a spliced evolutionary history. Furthermore, we discuss the obstacles of inferring the biogeographic history of the region.

2.2 Introduction

The Indian Ocean is the smallest, youngest and warmest of the three major oceans (Atlantic, Indian and Pacific). It bathes areas of great geologic and biogeographic interest,

including the African, Asian and Australian continents, the Indian subcontinent, the Indo-Australian Archipelago (IAA), the Arabic Peninsula, and important islands such as Madagascar, Seychelles, Comoros and the Mascarene Islands. Even though the Australian western oceanic crust has shown to be 140 million years (Ma) old – which suggests that the ocean’s initial opening may share that age (Heirtzler et al. 1973) –, the Indian Ocean’s origin is much more complex and recent, as it involves successive seas and oceans’ openings and closures (see below).

The Indian Ocean’s climate is affected by a monsoon system, especially in the Northern Hemisphere. The winds blow towards the southwest from October to April and in the opposite direction from May to October. That pattern leads to two very well delineated seasons – respectively, one cold and dry and another warm and rainy. The Southern Hemisphere is less influenced by the monsoon, with milder winds; nonetheless, it suffers with severe storms during summer due to the northeast winds (Clemens et al. 1991; Gadgil & Srinivasan 2011). In fact, the Indian Ocean’s monsoon is so marked that it can cause extensive changes in the oceanic currents of the north hemisphere: during the winter monsoon, two smaller gyres are formed at east and west of the Indian subcontinent (Tomczak & Godfrey 1994; Schott et al. 2002).

In addition to being the hottest ocean in the planet, the Indian Ocean also suffers with the highest levels of annual warming, especially in its western section. Such warming is intimately linked to global warming and to El Niño, and they show how El Niño can induce an abnormal warming on other oceans, which La Niña isn’t able to revert completely (Roxy et al. 2014). The abnormal oceanic warming could also be associated to the weakening of the rainfall regimen during the summer monsoon since precipitation weakening in the Indian Ocean is correlated to rises in the sea surface temperature (Roxy et al. 2015). With those peculiar warming properties, there is a blatant concern for the region in what touches global warming (Roxy et al. 2014; Latif et al. 2017; Aneesh & Sijikumar 2018).

Such outstanding recent findings underline how dynamic the geography of the Indian Ocean can be. It is not surprising that this ocean has become a prolific field for biogeography and a laboratory for studies on continental drift (see Hocutt 1987; Chatterjee et al. 2013). However, the region remains fertile to discoveries both about Earth’s crust dynamics (e.g. Bybee et al. 2010; Ashwal et al. 2017) and about the distribution of the organisms across the world (e.g. Pyron 2014; Jurado-Rivera et al. 2017).

2.3 Overview of the Indian Ocean's geology

The history of the Indian Ocean begins with Tethys, an ancient sea located in the eastern Pangaea. It lasted from the Devonian to the Cenozoic, but, it's now understood that Tethys had neither temporal nor spatial stability; in fact, it comprised several oceans that were opened and destroyed, leading to the current nomenclature: Paleo-Tethys (a sea that lasted from the Devonian to the Triassic), Meso-Tethys (from the Permian to the Cretaceous) and Ceno-Tethys (from the Triassic to the Cenozoic). The duration of these seas overlapped because the expansion of each of these oceans would eventually cause the collapse of the earlier ocean (Metcalf 2013). There is no semantic consensus on the origin date of the 'Indian Ocean' as we understand it. Some authors namely call 'Indian Ocean' the sea created at south of the Indian subcontinent by its movement northwards (e.g. Metcalfe 2013). Other authors only use the term Indian Ocean to refer to the ocean left after the collision between India and Eurasia at the transition Oligocene-Miocene, i.e., the ocean left after the complete destruction of Tethys (e.g. Hall 2012).

During the Late Jurassic, 160 Ma ago, Pangaea was breaking into two new continents, Laurasia and Gondwana. At this time, Gondwana was also initiating a fragmentation process that began by the separation of Australia and India. The Australia-India breakup forced the Ceno-Tethys northwards and started the collapse of the Meso-Tethys (Fig. 2.1a) (Chatterjee et al. 2013). The separation between India and Australia, 140 Ma ago, marked the beginning of the Cretaceous period (Fig. 2.1b) and, as India moved to the north, it carried Madagascar with it (Fig. 2.1c). The separation between these two landmasses only happened 90 Ma ago (Fig. 2.1d) (Raval & Veeraswamy 2003; Hall 2012; Metcalfe 2013) followed by the separation of the Seychelles sometime near the end of the Cretaceous (Collier et al. 2008). Meanwhile, Australia drifted eastwards; the rift between Australia and Antarctica was initiated by the Gondwana break-up but both continents were still widely connected until a second rift event in the Late Cretaceous. With the second rift event, Australia initiated the movement northwards (Hall & Keetley 2009). In the Palaeocene, when the IAA did not exist yet, the Sunda shelf was presented as a peninsula in south-eastern Asia (Fig. 2.1e). The IAA was originated from the movement of the Australian plate towards the shelf: throughout the Cenozoic, such motion generated thousands of volcanic islands, rotated Borneo anticlockwise and eventually set the current boundaries of the Pacific Ocean. The northernmost portion of the Australian plate, a promontory named Sula Spur, collided with the southernmost portion of the Sunda shelf in the

Miocene. The encounter caused the collision of several islands that gave rise to Sulawesi. Sulawesi's formation established the contact between the Australian and Eurasian plates and ended the wide connection between the Pacific and Indian oceans (Hall 2011, 2012).

During the Cenozoic, the Indian Ocean's currents were unstable due to India's movement (Barron & Peterson 1991). However, after its collision with Asia, the shallow circulation became simpler, basically resumed to an equatorial flow off the Pacific towards Africa and a flow back to the Pacific from the southwest (Barron & Peterson 1991; Thomas et al. 2003). At 14 Ma, the strait between Sula Spur and the Sunda shelf became so narrow that the equatorial flow back to the Pacific was interrupted and a strong westward equatorial current took place in the Indian Ocean (Fig. 2.1f) (Gourlan et al. 2008). The current, called Miocene Indian Ocean Equatorial Jet (MIOJet), lasted for over 10 Ma and not only increased the transportation westwards but also diminished the transportation eastwards. The establishment of the MIOJet altered that dynamics in such a way that the surface circulation became almost restricted to the jet (Barron & Peterson 1991; Gourlan et al. 2008).

The complexity of the Indian Ocean's evolution drew attention of geologists for decades and now it is very well understood. Such complexity is mostly due to the India's movement: as the subcontinent passed to the north, it influenced not only the oceanic circulation but also climate and the shape of the continents as multiple volcanic hotspots became active. The most significant hotspots in the Indian Ocean are the Kerguelen hotspot and the Réunion hotspot. The Kerguelen hotspot began to erupt in the Early Cretaceous and is therefore related to the Antarctica-Australia-India break-up (Coffin et al. 2002). The Réunion hotspot (and the Deccan Traps) became active in the K-Pg boundary and may be related to the India-Seychelles separation. Each of these hotspots produced a long ridge across the Indian Ocean – east and west of India respectively (Fig. 2.1f) – with subaerial elements in the Cretaceous and Cenozoic (reviewed by Chatterjee et al. 2013). The Deccan Traps are particularly interesting as, for decades, their eruption has been suggested to be an explanation the mass extinction occurred at the end of the Cretaceous period (e.g. McLean 1978; Venkatesan et al. 1993; Keller et al. 2011), competing with the popular bolide impact hypothesis, which itself presents compelling evidence (Schulte et al. 2010). Some authors, however, don't see the two hypotheses as alternatives (Schoene et al. 2014; see Keller 2014) and some studies have suggested causal connectivity between the two, where the impact would have set off the tectonic instability (Renne et al. 2015).

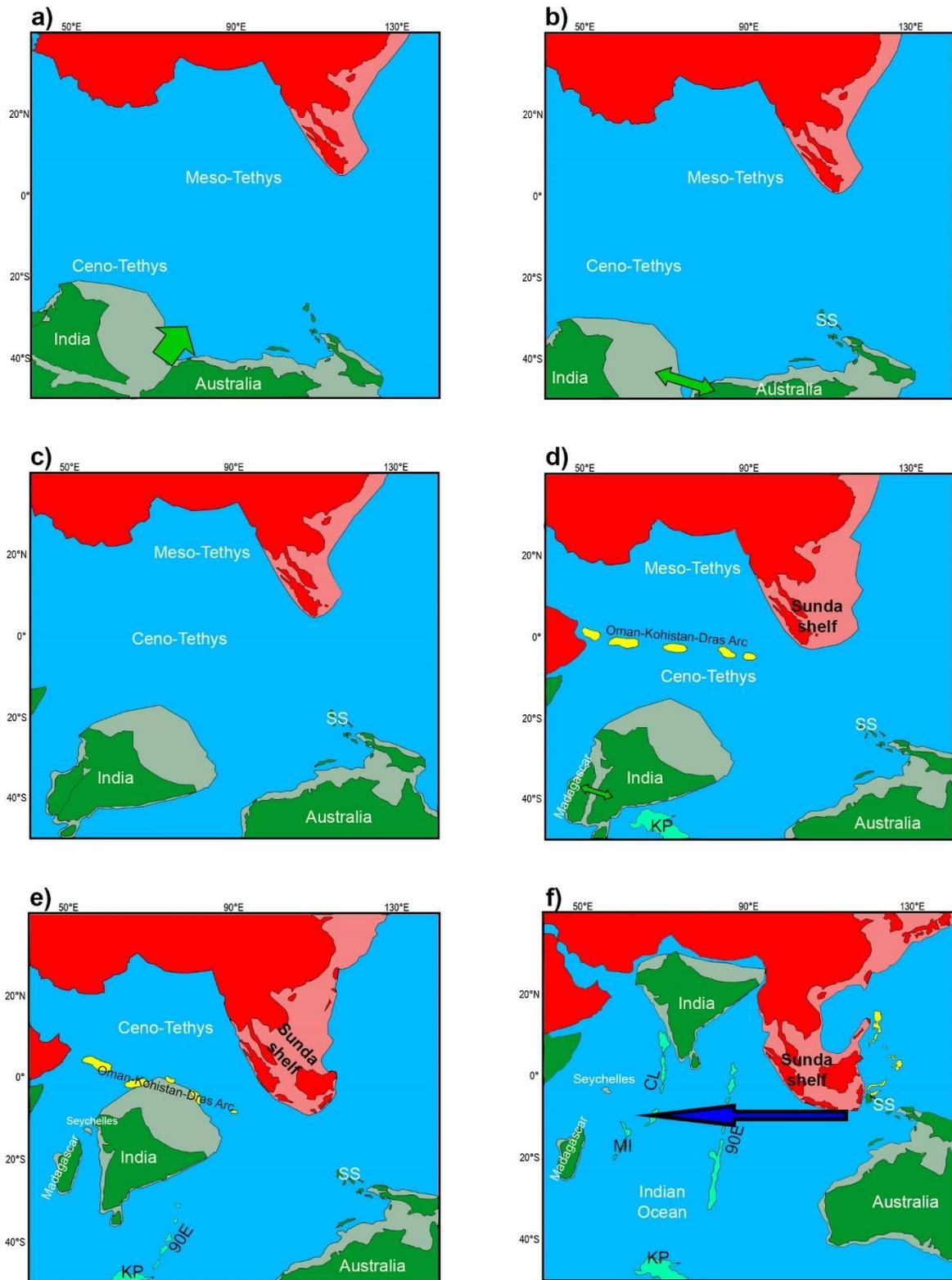


Figure 2.1: Geologic reconstruction of the Indian Ocean's continents through the Mesozoic and Cenozoic at 155 Ma (a), 140 Ma (b), 110 (c), 90 Ma (d), 65 Ma (e) and 14 Ma (f). Laurasian landmasses are represented in red and Gondwanan landmasses are represented in green. Areas in yellow are volcanic islands and in cyan are volcanic ridges (90E: Ninetyeast Ridge; CL: Chagos-Laccadive Ridge) and plateaus. The oceans are in blue. The lightly coloured areas next to the continents represent the respective continental shelves, which may or may not be emerged depending on the time. Dashed lines represent boundaries of continental break-ups; green arrows represent major break-up events; and the blue arrow represents the MIOJet. KP: Kerguelen Plateau; MI: Mascarene Islands; SS: Sula Spur. Modified from Hall 2012 after Chatterjee and Scotese 2010 and Chatterjee et al. 2013.

2.3.1 On the Eastern Indian Ocean

The most obvious outcome of the collision between the Indian subcontinent and the Eurasia 50 Ma ago was the formation of the Himalaya. The uplift of the Himalaya had an extensive influence over the climate, not only Asian but worldwide: it is now seen as a major contributor to the establishment of the monsoon system and it favoured the glaciation of the Northern Hemisphere (reviewed by Chatterjee et al. 2013). Nevertheless, the encounter also helped to push the Sunda shelf eastwards, and thus contributed to the formation of the IAA and to the rotation of the Greater Sundas – Borneo, Sumatra and Java (Allen et al. 2008). By the time of that collision, Australia, still connected to Antarctica, was accelerating its northwards movement, triggering the process of emergence of the IAA as described above (Hall 2002). The IAA formation defined the eastern boundary of the Indian Ocean and changed its dynamics completely.

As said in the previous section, though, the IAA did not exist in the Early Cenozoic and the Sunda shelf was a promontory at the south-eastern portion of Eurasia. During the Palaeocene, when the climate was similar to nowadays (perhaps warmer and wetter), the Sunda shelf probably exhibited environmental complexity, with mountains, rivers and equatorial forests. However, the climate changed in the Middle Miocene, leaving the area drier and more seasonal (Morley 2000). The emerged area decreased consistently until the Late Miocene, when the contact between Sula Spur and the Sunda shelf initiated in the Early Miocene expanded and the Makassar strait (a rift currently between Borneo and Sulawesi) was formed (Hall 1996, 2002). The encounter between Sula Spur and Sunda also caused the uprising of mountains and landmasses such as Borneo (Hall 2002) (see fig. 2.1e). The uprising of mountains, the expansion of shallow seas and the changes in the circulation caused by the MIOJet contributed to the establishment of a wetter climate, which made the forests expand again in the region (Morley 2000). The collision also augmented the emerged area of the Wallacea, the middlemost region of the IAA, including Sulawesi, Halmahera, Taliabu and the Lesser Sundas (Timor, Flores, Sumba etc.) (Hall 2002).

2.3.2 On the Western Indian Ocean

The breakup between India and Madagascar in the Cretaceous initiated the formation of the Western Indian Ocean's islands. The fragmentation is important to geology in multiple levels; more recently, new evidence culminated in the discovery of a previously unknown

microcontinent called Mauritia. These findings reveal that the India-Madagascar separation also involved the fragmentation of the cited microcontinent (Torsvik et al. 2013; Ashwal et al. 2017). Soon after the separation of Madagascar, at some point in the Late Cretaceous, the Indian subcontinent broke from the Mascarene Plateau, where the Seychelles are located. The Mascarene Plateau is a submarine plateau of mixed origin: whereas the northern part (with the Seychelles) is granitic, the southern part, which includes the Mascarene Islands is volcanic (Ashalatha et al. 1991). The granitic portion is actually a 750 Ma old microcontinent whose fragmentation began in the Carboniferous and may have triggered the India-Madagascar-Seychelles separation (Plummer & Belle 1995; Torsvik et al. 2001). The Seychelles are therefore the oldest known granitic islands, which implies that their geological stability is considerable.

As said above, the Réunion hotspot became active in the K-Pg boundary and was responsible for the formation of the southern part of the Mascarene Plateau. The hotspot kept active for millions of years and generated many islands as well as the Chagos-Laccadive ridge, south-eastern to India (Fig. 2.1f). The Laccadive Islands, the Maldives and the Chagos Archipelago were the first islands to emerge between 60 and 45 Ma ago. The Mascarene Islands would only begin their slow emersion at 35 Ma: first, the Saya de Malha bank emerged followed by the Nazareth Bank, both submarine nowadays; Mauritius would emerge next, approximately 10 Ma ago; Réunion and Rodrigues were the last ones at 2 Ma (Ashalatha et al. 1991; Verzhbitsky 2003). As for Madagascar, after its breakup from India, the island moved away from Africa very slowly. In part, it was due to the Réunion hotspot's activity and to the expansion of the Chagos-Laccadive ridge across the Mascarene Plateau (Verzhbitsky 2003; Ali & Aitchison 2008). Even though it is probably not related to this movement, the origin for the Comoros hotspot (between Africa and Madagascar) is still a controversial topic. The Comoros Islands (including Mayotte) are an outcome of this hotspot; the region still presents a strong volcanic activity and the islands are somewhat isolated from both Africa and Madagascar (Esson et al. 1970; see Barber-James 2008).

2.4 The Indian Ocean's biogeography

‘Probably no theory has shaken the foundations of biogeography more than Croizat's (1958) *panbiogeography* which broke from traditional dispersal models (...) in favour of vicariance theory.’

Charles H. Hocutt (1987), *Evolution of the Indian Ocean and the drift of India: A vicariant event*

In the previous sections, it has been made clear how complex the history of the Indian Ocean is. Such complexity can be translated in richness when it comes to biodiversity. The Indian Ocean is intimately related to the history of five continents (Africa, Antarctica, Australia, India and Eurasia); therefore, after the emergence of the vicariant biogeography, it became a laboratory for studies in historical biogeography (see Hocutt 1987). Such studies suggest not only major vicariant hypotheses involving plate tectonics but also draw attention to one of the new directions given to historical biogeography by the vicariant theory: ‘areas of endemism’ (Nelson & Platnick 1981). The concept and the criteria to define areas of endemism has been largely debated: Henderson (1991) criticise the very study of the relationship of areas in biogeography and pointed that the criteria to identify areas of endemism are poorly defined. Linder (2001), however, argues that methodological refinement can make the delimitation of areas of endemism objective. Following the notion that the concept of areas of endemism involves only space and not time, Noguera-urbano (2016) proposes that studies on endemism would be enriched by the evaluation of the relations between endemic species through time, i.e., through phylogenetic information. In spite of such heated debate, it is noteworthy that some of the most iconic areas of endemism are bathed by the Indian Ocean: Australia, Madagascar, the Mascarene islands, the Seychelles, just to cite a few.

India itself has been seen as a good model for studies in historic biogeography for the time it spent travelling in isolation across multiple climatic zones. India’s biogeographic isolation has been questioned for decades: Briggs (1989) points out that, even though the geophysical evidences indicate 100 Ma of an isolated subcontinent, the palaeontological data does not exhibit the expected endemism for such scenario. However, Patterson and Owen (1991) show that some of Briggs’ arguments fail to support his claims of a non-isolated India. Over two decades later, Verma et al. (2016) reviewed the prolific palaeontological discoveries since the 1980’s in the region and showed that some of the endemism expected for the Late Cretaceous period was in fact there, but mixed elements (Gondwanan and Laurasian) indeed predominate in the fossil assemblages. An intriguing fact of such debate is that the current taxa represent the isolation indicated by the geological history better than (sometimes even in disagreement with) the fossil assemblages, with some studies providing evidence for an ‘Out-of-India’ hypothesis for the presence of Gondwanan forms in Asia (e.g. Conti et al. 2002; Gower et al. 2002; Biju & Bossuyt 2003; Datta-Roy & Karanth 2009; Whatley 2012). Many

bridges have been hypothesised to explain the presence of taxa of foreign affinities in India. Chatterjee and Scotese (2010; also see Chatterjee et al. 2013) present compelling evidence for two of those bridges: the Oman-Kohistan-Dras Arc would serve as a connection to Africa in the Late Cretaceous (and to Asia in the Cenozoic) (see fig. 2.1d,e); at the same time, the Ninetyeast Ridge plus the Kerguelen Plateau would connect Indian and South American faunas via Antarctica (see fig. 2.1e,f). Other proposed bridges would have linked Africa and India via Madagascar through currently submerged elements of the Mascarene Plateau and the Chagos-Laccadive ridge (e.g. Sahni 1984) or through the Seychelles (Rage 1996, 2003; Ali & Aitchison 2008).

Until the end of the Cretaceous, India carried Madagascar and the Seychelles in its journey northwards. Madagascar has one of the highest levels of endemism in the world, including 100% of the mammals and amphibians, 92% of the reptiles and over 90% of the plants (reviewed by Vences et al. 2009). In addition to this outstanding endemism, Madagascar has a very heterogeneous landscape, similar to larger continents, with several different climatic and vegetation zones (de Wit 2003), a scenario that drew the attention of Alfred Russel Wallace (1881) himself: ‘Madagascar possesses an exceedingly rich and beautiful fauna and flora, rivalling in some groups most tropical countries of equal extent, and even when poor in species, of surpassing interest from the singularity, the isolation, or the beauty of its forms of life.’ Its fauna has always been very Gondwanan, including derived Abelisaurid dinosaurs similar to those found in South America in the Late Cretaceous (Sampson et al. 1998), but Laurasian forms invaded the island very soon in the Tertiary, including reptiles, mammals and plants (Rage 1996; Schatz 1996; see Krause 2010). It is geologically understood that Madagascar became isolated as soon as it was detached from India (Reeves & de Wit 2000; see Fig. 2.1d) and such scenario, as it is for India, presents opportunities for the proposition of new land bridges. One of the most debated (and no longer accepted) bridges for Madagascar is the so-called ‘Lemuria’ proposed by van Steenis (1962) for the Cretaceous. Schatz (1996) proposed that, though Lemuria is not supported by any geological evidence, the Eocene India and elements of the Mascarene Plateau could have served as “Lemurian Stepping-stones” for Laurasian organisms. Since then, several studies have evoked these stepping-stones as an explanation for faunal/floral exchanges between Laurasia and the Gondwanan elements of the Indian Ocean (e.g. Wikström et al. 2004; Klaus et al. 2006; Trénel et al. 2007; Moyle et al. 2012).

Only in 2010, Warren et al. have provided strong evidence for the existence of such stepping-stones across the Seychelles and Mascarenes. These archipelagos, though smaller in size and poorer in ecological complexity when compared to Madagascar, exhibit remarkable levels of endemism: for example, in both archipelagos over 50% of the native resident species are endemic (Adler 1994). Separated from India in the Late Cretaceous, the Seychelles are much older than the Mascarenes (see previous section); yet there are examples of taxa that can be found in both archipelagos, but are not found neither in Madagascar nor in Africa which is in agreement with the stepping-stone hypothesis (Warren et al. 2010). Also in agreement with an island presentation that favours the Lemurian Stepping-stones is the historic distribution of tortoises (Braithwaite 2016; Cheke et al. 2016; Hansen et al. 2016; Hawlitschek et al. 2016) and other reptiles (reviewed by Hawlitschek et al. 2016) in the Indian Ocean. Nevertheless, Thornton et al. (2002) questions the relevance of stepping-stones in the colonisation of islands. As he discusses, since good dispersers will be able to reach a location without the aid of a stepping-stone whereas poor dispersers may fail to do so even in the presence of a stepping-stone, the distance covered by the stepping-stone should be near the dispersal limit of a said species. Moreover, environmental divergences between the target island and the stepping-stone may play an important role in defining the spectrum of species that are able to use such dispersal path. To date an island's endemism is also a difficulty to study and hypothesise on the historic dispersal across islands: Pillon and Buerki (2017) showed that extinction has been an overlooked phenomenon in biogeography and insular lineages may have diverged previous to the emergence of the island itself.

On what comes to marine organisms, things may seem a little simpler, but they also failed to fulfil historic biogeography's predictions; Heads (2005), reviewing the prominence of panbiogeographic patterns in marine biota, pointed to many works that demonstrate a vicariant origin of the marine species from Tethyan ancestors. In fact, some interesting cases show a clear endemism boundary between eastern and western Indian Ocean: for instance, both extant species of coelacanths are found in the Indian Ocean, one in Comoros and the other in the IAA (Holder et al. 1999; Springer 1999). Other examples of such east-west provinces are not rare and can be observed in the blenniid genus *Ecsenius* McCulloch, 1923 (Springer 1988), in pigeons – where several genera found in the Malagasy islands exhibit transoceanic relationships, including *Alectroenas-Drepanoptila*, *Raphus/Pezophaps-Caloenas* and multiple species in the genus *Treron* Vieillot, 1816 (Shapiro et al. 2002; Pereira et al. 2007; Gibb & Penny 2010) –, and in lizards (e.g. Rocha et al. 2006; Heinicke et al. 2014) to cite a few. Some

of these, however, may not represent Croizat's school so well as some may believe: developments in dating speciation events in a phylogeny using molecular data made possible to discard tectonics as responsible for some disjunct distribution across the taxa (reviewed by de Queiroz 2005). For example, there are some single species cases of this eastern-western division in the Indian Ocean – e.g. the gastropod *Cerithium gloriosum* Houbriek, 1992, found in the Comoros and in Borneo (Houbriek 1992), and the freshwater shrimp *Caridina typus* H. Milne-Edwards, 1837, a species complex found across all of the Indian Ocean (Bernardes et al. 2017 and Chapter 3 below) – and imply that gene flow is still happening or has happened in recent times.

The Indian Ocean bathes areas intensively targeted by the panbiogeography school, but still holds many puzzles unanswered. Linder (2001) criticises the approach by some panbiogeographers to demonstrate vicariance by evaluating incongruence in the composite areas: '(...) this means that incongruence could be caused by either poor area delimitation or dispersal, with no mechanism for deciding between these two hypotheses. The possibility of refuting process explanations of vicariance is thus weakened' (p. 893). The methodological development in the last two decades should have eliminated such arguments from authority used to criticise new findings that are incongruent with the previous literature, but the debate stands and much of Indian Ocean's biogeographical history remains unsolved. Here we shall review the molecular phylogenetic and phylogeographic works made in the Indian Ocean and apply the latest methods to them in order to assess their biogeographic information. This will be then used to propose a history to the Indian Ocean continents and islands.

2.5 Materials and methods

2.5.1 Studies chosen

Search for studies to be used in this meta-analysis was conducted in Google Scholar and Web of Science. Our choices were based on three criteria: (1) they should be centred in either terrestrial or freshwater taxa; (2) they should be based on molecular data, namely DNA sequences; and (3) they should be focused on phylogenetic reconstructions. With such criteria, we used seven published researched about the landmasses in the Indian Ocean (Table 2.1). We chose papers that included more than one of the defined landmasses: continental Africa, continental Asia, Australia, Comoros, India, Indo-Australian Archipelago, Madagascar, Mascarene Islands and Seychelles.

Table 2.1: Works analysed in this dissertation

Taxon	Taxonomic level	Distribution	Reference	Calibration	Markers
Pteropodidae	Family	Af, As, Au, C, In, IAA, M, MI, S	Almeida et al. 2011, 2014	<i>Hassianycteris kumari</i> Smith et al. 2007; Dates from Agnarsson et al. (2011)	12S, 16S, CytB BRCA1, RAG1, RAG2, vWF
Chamaeleonidae	Family	Af, As, C, In, S	Raxworthy et al. 2002 Tolley et al. 2013	Same as in Tolley et al. (2013)	16S, ND2, ND4 PRLR, RAG1, CMOS, AKAP9, BACH1, BACH2, BDNF, MSH6, NKTR, REV3L
<i>Cryptoblepharus</i>	Genus	Af, Au, C, M, MI	Rocha et al. 2006	Rates for 12S and 16S after Skinner et al. (2007)	12S, 16S
<i>Ebenavia inunguis</i>	Species	Af, C, M, MI	Hawlitshcek et al. 2017	Rates for 12S, CytB and RAG2 after Carranza & Arnold (2012)	12S, CytB, COI RAG2, PRLR
Boidae	Family	Af, As, Au, M, I	Noonan & Chippindale 2006	Same as in Noonan & Chippindale (2006)	CytB RAG1, BDNF, CMOS, NT-3, ODC
Anatini	Tribe	Af, As, Au, I, IAA, M	Mitchell et al. 2014	Same as in Mitchell et al. (2014)	12S, CytB, COI, ND2
Neobatrachia	Suborder	Af, As, Au, I, IAA, M, S	Frazaõ et al. 2015 Feng et al. 2017	Same as Feng et al. 2017	12S, 16S, ND1, CytB C-MYC2, C-MYC3, H3A, POMC, RHOD, RAG1, SIA, TYRPREC

Area abbreviations: Af - Africa, As - Asia, Au - Australia, C - Comoros, I - India, M - Madagascar, MI - Mascarene Is., S – Seychelles

2.5.2 Phylogenetic analyses

All phylogenetic analyses were conducted in BEAST 2 v. 2.4.0 (Bouckaert et al. 2014) and aimed to replicate the results obtained by the original authors in order to test them – if they could not be replicated they were either mistaken or not clear about the used methodology. However, we observed three issues with repeating the exact same procedure described by the authors. First, some of the works do not give enough detail to allow a precise replication of

their work. Second, many works fail to apply the best tools or the appropriate empirical priors to their Bayesian analyses. Third, to facilitate the obtention of convergence was a main concern for us due to the large number of Bayesian analyses to be performed. With these issues in mind, we established a routine to set each analysis:

(I) Data handling: we would use every information made available by the authors in our first approach to the data (alignment, partitioning, priors etc.). Nonetheless, some of the works did not use the best approach to treat the data before the analysis: for example, many papers used ribosomal markers without applying a secondary structure to their alignment (without using a secondary structure as a base for the alignment of such markers, the proposed primary homology will inevitably present errors);

(II) Substitution model: jModelTest 2.1 (Darriba et al. 2012) was used to calculate the best named model (e.g. JC, HKY, GTR) for each marker or partition indicated by the authors, but all analyses were performed using a reversible-jump method to choose the best substitution model (Huelsenbeck et al. 2004; Bouckaert et al. 2013). The reversible-jump method was chosen instead of a named model because (1) it facilitates convergence since it is able to find the least parametrised model to fit the data (see Grummer et al. 2014) and (2) it makes the phylogenetic estimation independent of a specific model by integrating it over the parameters' uncertainty (Huelsenbeck et al. 2004). However, our previous experience shows that a better convergence is obtained when offset values for the model parameters (such as invariant proportion) are empirically obtained through jModelTest;

(III) Molecular clock: all molecular clocks were initially set to relaxed with a lognormal distribution. The strict clock is a special case of the relaxed clock where there is no variation of the rates along the branches. After the preliminary analyses, those markers that presented the distribution of the clock's standard deviation centred in zero were set to a strict clock. The remaining relaxed clocks would then be tested for an exponential distribution. If there were differences in convergence or topology, the two distributions, lognormal and exponential, would be tested against each other through an AICM approach (Raftery et al. 2007; but see Baele et al. 2012);

(IV) Tree priors: all trees were initially set to either a Yule or a Coalescent model and lognormal distribution for all priors. A Yule model would be set for multispecies analyses and Coalescent models would be set for analyses involving a single species. Yule and Birth and

Death models were alternatives according to the convergence. The distribution and parameter values of the priors were set according to assessments made through TRACER v1.6 (Drummond et al. 2012);

(V) Calibration: as one of the main concerns of this study is to describe the history of the taxa distribution, we were very attentive to calibrations. Since some of the works were not preoccupied with dating their phylogeny, we sought in the respective taxon literature for available calibration tools. Calibration methods of three types were found across the works we found – fossil, molecular rate and secondary calibrations obtained from larger phylogenetic studies –, and, therefore, we saw no reason to avoid any of these methods. Geological calibrations such as island emergence were not used since we agree with the points raised by Pillon and Buerki about the relationship between a taxon age and an island age (2017). We used the calibration points and substitution rates of the source paper (when available) and/or from secondary sources (see Table 2.1). All of our analyses were calibrated.

Alignments were done separately and, due to the variety of markers, several methods and softwares were used. All analyses were run for 500 million generations sampling every 50 000 generations. They were visually assessed in TRACER and submitted to either 20% or 25% burn-in, depending on the convergence obtained. Final maximum clade credibility trees were obtained through *TreeAnnotator* by using the same burn-in percentage. All analyses in BEAST2, MAFFT (Kato & Standley 2013) and jModelTest were run in CIPRES (Miller et al. 2010).

Table 2.2: Coordinates used for areas of interest in BayArea

Area	Latitude	Longitude
Africa	-1.2500	25.7225
America	10.7769	-79.1874
Asia	39.7054	81.8409
Australia	-24.1385	133.3899
Comoros	-12.1837	44.0731
Europe/ Mediterranean	36.637	-2.9501
IAA	1.2500	113.4847
India	20.6835	79.1195
Madagascar	-19.4258	46.6878
Mascarene	-20.2856	57.5659
New Zealand	-43.6889	171.002713
Pacific	-13.9791	-173.2304
Seychelles	-4.7005	55.4899

2.5.3 Biogeographic analyses

We inferred ancestral areas of distribution in RASP 4.0 (Yu et al. 2015) – RASP 4.0 also implements BioGeoBEARS (Matzke 2013) and thus allows to also use all features available at this package. Localities of the specimens and distribution of species were obtained from the original papers or from web sources (IUCN, The Reptile Database, GenBank). These localities were designated to one or more of each areas described in table 2.1 – note that IAA includes part of the continental Southeast Asia since they were connected by land through much of the Cenozoic. Models were tested through BioGeoBEARS, but both DEC (Ree & Smith 2008) and BayArea (Landis et al. 2013) as implemented in both RASP and BioGeoBEARS were run in order to find incongruences between methods. BayArea was run both with and without coordinates (table 2.1) for the same reason. All analyses were run for 20 million generations sampling every 4 000, with unconstrained settings for dispersal with a maximum range size of three areas.

BayArea/DEC inferred areas were represented on the trees obtained in BEAST; inferred ancestral distributions with probabilities under 0.5 were removed from further consideration. Also removed from consideration were dispersal/vicariance events that did not involve the Indian Ocean (e.g. Africa-America or America-Australia).

2.6 Results

2.6.1 Ancient taxa diversification

Figure 2.2 shows the biogeographical history of each taxon. On the left are the time-calibrated trees with the calculated ancestral areas of distribution. Each colour on the branches represents an area according to the legend (and to the map) and more than one colour (striped branches) represent more than one area included in the distribution range (to avoid confusion, the range is described at the branch). Time table is located below the phylogenies for temporal orientation based on the mean age of each node. On the right are the maps with the respective described events. Arrows are coloured based on the Period/Epoch in which the event occurred.

The meta-analyses of the published data identified a very complex biogeographic history mostly based on dispersal throughout the Cenozoic. Only three events were identified

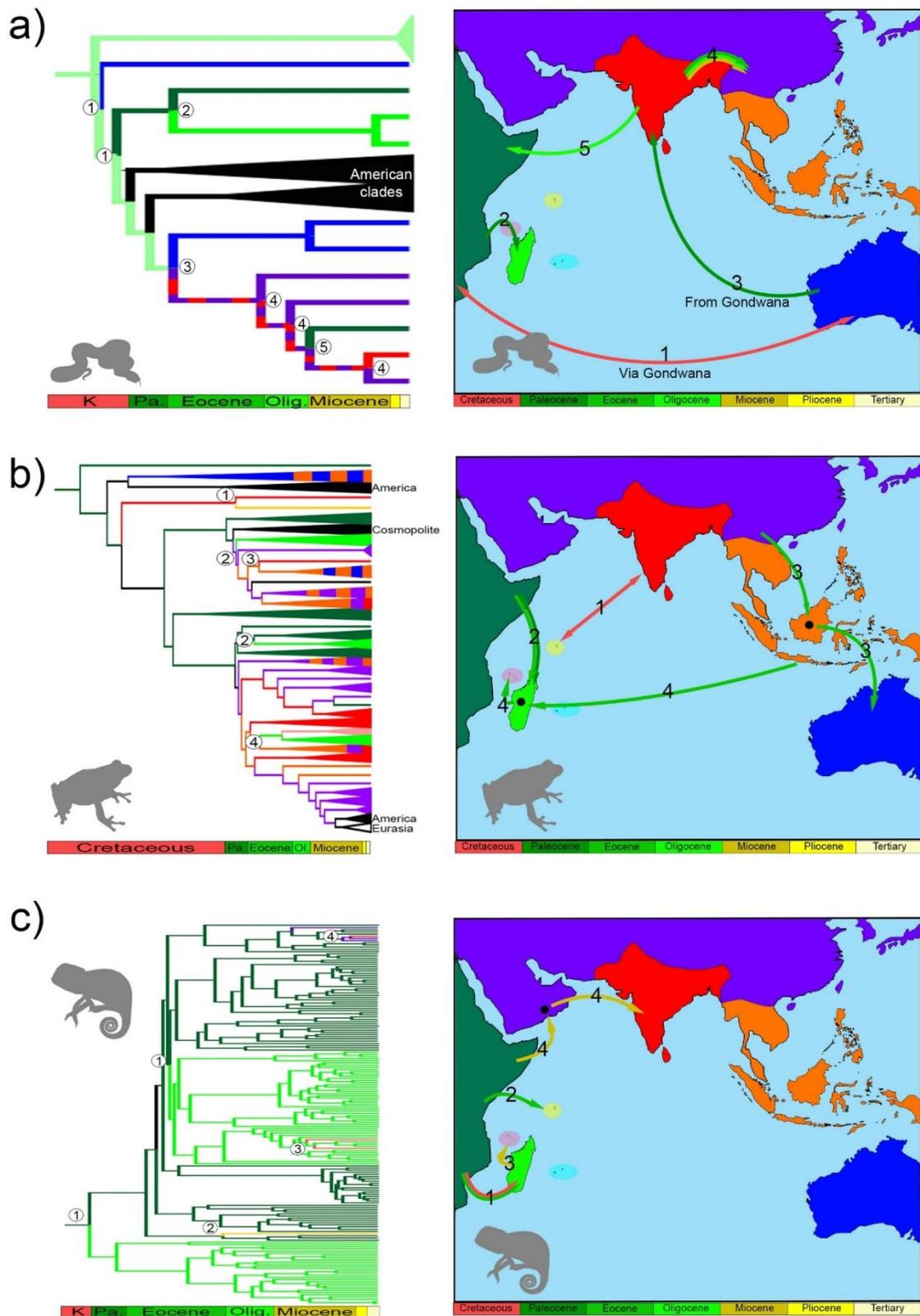


Figure 2.2: Ancestral area reconstruction and biogeographic history. On the left are the ancestral reconstructions as inferred by BayArea and DEC on calibrated phylogenies for Boidae (a), Neobatrachia (b), Chamaeleonidae (c), Pteropodidae (d), Anatini (e), *Cryptoblepharus* (f) and *Ebenavia inunguis* (g). The branches colours represent the correspondent area in the maps on the right according to the legend. The maps on the right represent our interpretation of each taxon's biogeographic history. Areas not represented in the map are presented at the branches. The arrows represent biogeographic events and are coloured according to the Period/Epoch in which they happened according to the legend below each map. One-way arrows represent dispersal and two-way arrows represent vicariance.

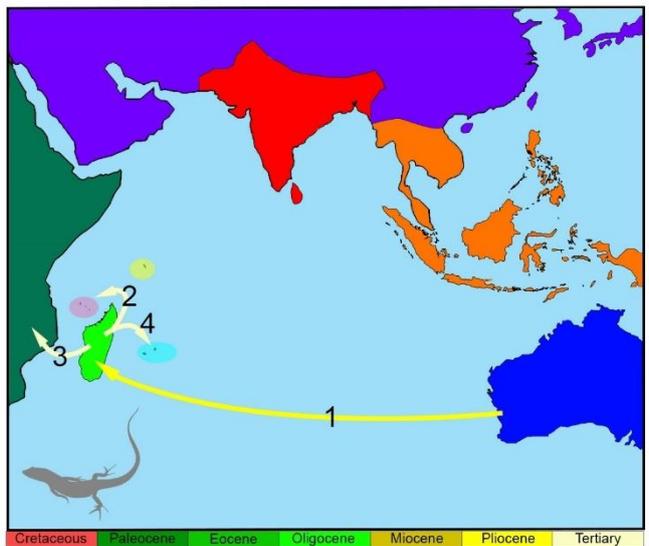
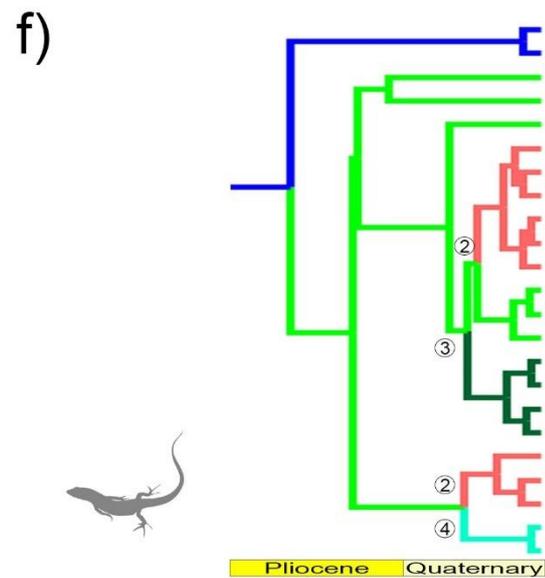
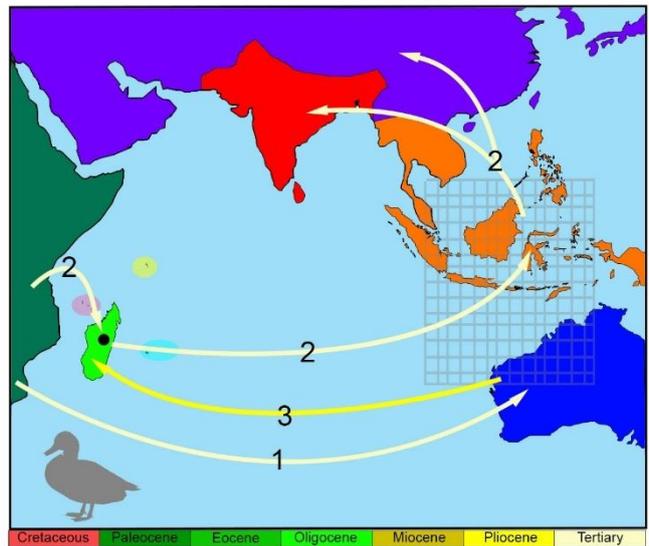
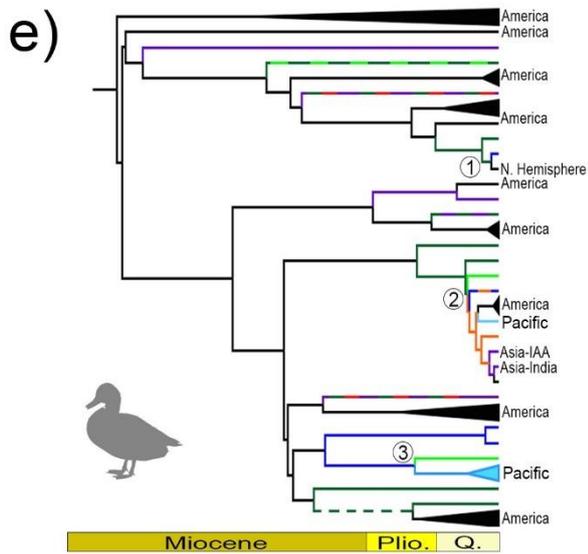
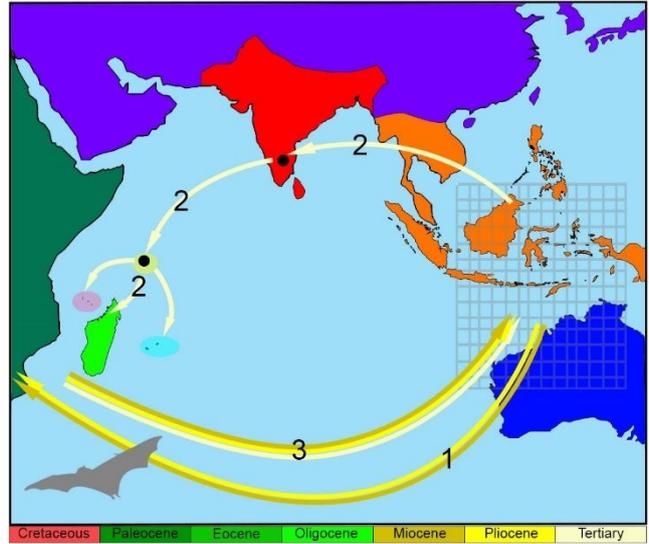
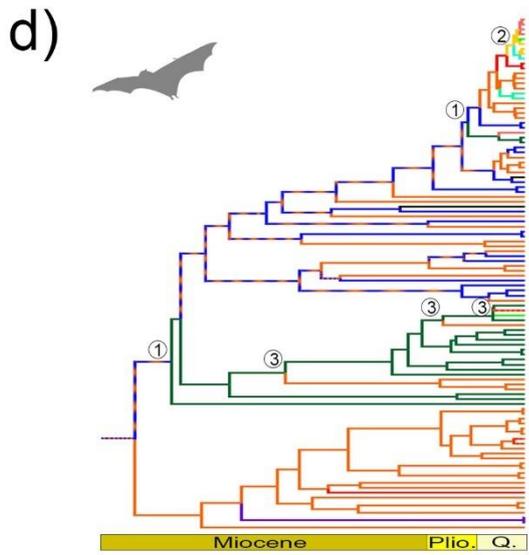


Figure 2.2: Cont.

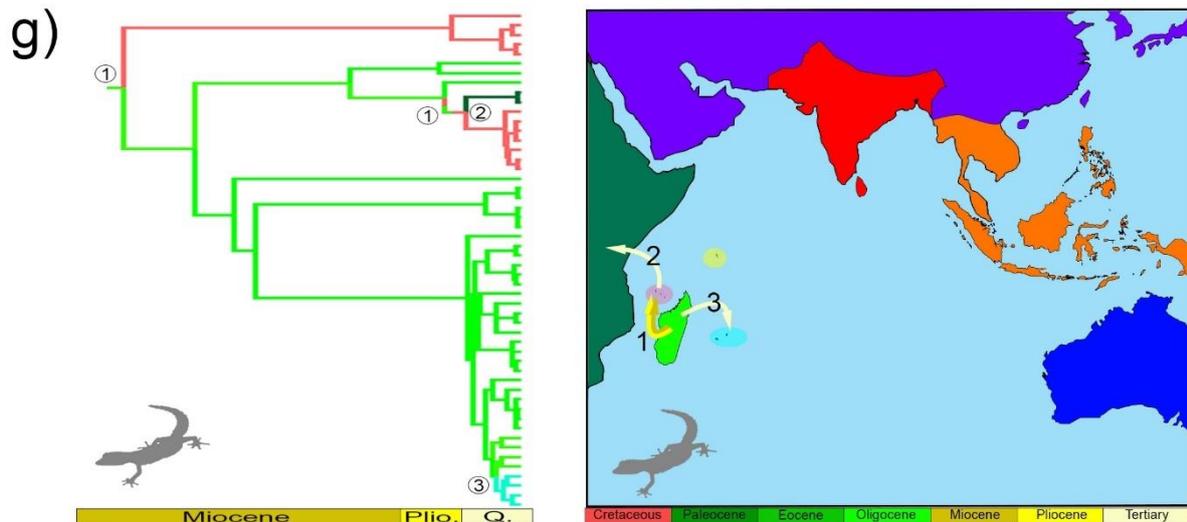


Figure 2.2: Cont.

to take place in the Mesozoic (Fig. 2.2a,b,c) and not coincidentally among those are also the only two vicariance events that we were able to point out. Boid snakes, chameleons and frogs are part of a non-volant less vagile herpetofauna (with some remarkable exceptions among the Neobathachia). However, all three taxa are very old in age and should be good models to assess vicariance events and ancient continental history. For instance, Boids were widely distributed across Gondwana and their current distribution seems to be intimately related to its fragmentation. Nonetheless, their colonisation of India seems to have been caused by dispersal from Antarctica-Australia (Fig. 2.2a). Alternatively, Gondwanan frogs present an example of an Indian clade that travelled with the subcontinent and went through a vicariance speciation with the separation of the Seychelles (Fig. 2.2b).

Although complex histories are offered by boids and frogs, chameleons exhibit a much simpler scenario: they remained a typically African taxon with multiple dispersals to Madagascar. Though the first dispersal to Madagascar occurred by the end of the Cretaceous, their diversification happened entirely through the Cenozoic.

2.6.2 Volant taxa diversification

Figure 2.2 (d) and (e) show respectively the biogeographic history of the pteropodid bats and dabbling ducks, both volant highly vagile taxa. The pteropodid bats seem to have been originated in the Australia-IAA and diversified from this area, including several dispersal events between this area and Africa. At the time of such dispersal toward Africa, India was already in contact with Eurasia and could not serve as a stepping-stone – and this is also indicated by the IAA origin of the Indian species – but strong westwards surface currents were

present at the time in the Indian Ocean (Gourlan et al. 2008). Winds are intrinsically related to surface currents and, as some pteropodid species have high migratory capabilities (Fenton 2001), they could have either flied or used rafts to travel across the ocean.

The same transoceanic dispersal occurred several times for the Anatini, but in much more recent times (late Pliocene and Quaternary). These animals are known to be migratory and their range cover all continents bar Antarctica, thus it is safe to assume that such dispersal was accomplished simply by flight.

2.6.4 Recent taxa diversification

As expected, the histories of taxa in genus (*Cryptoblepharus* Wiegmann, 1834; Fig. 2.2f) and species (*Ebenavia inunguis* Boettger, 1878; Fig. 2.2g) levels involve much more recent events. The distribution of both taxa is centred in Madagascar and both were able to colonise Africa and the Comoros and Mascarene islands. *Ebenavia inunguis* was apparently able to use the Comoros as stepping-stones to reach Africa. These two lizard taxa present similar dispersive ability and the dispersal events inferred by ancestral area reconstruction seem to be inside their reachable area.

2.7 Discussion

2.7.1 On an isolated India and its relationship with Seychelles and Madagascar

The Indian subcontinent (and its related islands, Madagascar and Seychelles) travelled through a very dynamic region with the emergence and termination of several landmasses. The idea that it was completely isolated is thus naïve. We were able to recover dispersal events to India from Gondwana, Eurasia and Africa-Arabia throughout the Cenozoic whereas only the Neobatrachia show a clade being originated by India's initial break-up. India, as reviewed by Verma et al. (2016), have its own endemism, but its fauna was composed by mixed elements already in the Cretaceous. The fact that Indian Cretaceous fauna is shared by Madagascar, Africa and South America suggests that the subcontinent probably remained connected with Gondwana at the beginning of its separation. In addition, the early acquisition of Laurasian forms in the fossil assemblage indicates that some connection should exist in the north as well. The problem, then, lies in identifying when and how these migrations happened. Chatterjee and Scotese (2010) gave a considerable insight into applying the geological knowledge into this matter, but the dates of their proposed bridges are still a matter of debate. Furthermore,

some groups in our meta-analyses still present evidence for transoceanic recent dispersal that cannot rely upon ancient bridges (Fig. 2.2b,c,d,e). That may be the reason why recent and contemporary taxa represent India's endemism better than the fossil species. Nevertheless, before drawing conclusions, we must put more effort into including Indian taxa to the meta-analyses as these recent transoceanic dispersals are mostly from volant organisms.

Among the bats of the genus *Pteropus* Brisson, 1762 a recent quick diversification shows an interesting scenario of dispersal. A group from the IAA, colonised India, then, Seychelles and, from there, Madagascar, the Comoros and Mascarene islands (Fig. 2.2d). For the chameleons, a much simpler situation is seen when it comes to Indian clades: they used the Arabic peninsula as a stepping-stone to reach India (Fig. 2.2c). The ancestral area reconstruction of Gondwanan frogs indicates that Asian clades were mostly generated by dispersal from Asia and IAA with one notable exception (*Nasikabatrachus* Biju & Bossuyt, 2003) that originated the Seychellois genus *Sooglossus* Boulenger, 1906 by vicariance at the K-Pg boundary (Fig. 2.2b). Boids, a relatively ancient taxon, show an early dispersal to an already drifting India in the Paleocene, perhaps through the Kerguelen Plateau-Ninetyeast Ridge (Chatterjee & Scotese 2010). India would then have served as a raft to allow Asian colonisation (Fig. 2.2.a).

These scenarios show heterogeneity in the colonisation of India. As discussed in the introduction, there is plenty of evidence for India serving as a Gondwanan ark toward Laurasia. Conversely, there's also plenty of evidence for early invasion by Laurasian forms (e.g. Prasad 2012; Khosla 2014). At the end of the Cretaceous, the Indian Ocean was a very troublesome region, with intense volcanism, breaking continents and changing climate; those times may hide scenarios that we are currently unable to recover. However, our analyses show that the influence that dispersal has exerted on India's colonisation should not be ignored.

2.7.2 On the colonisation of the Indian Ocean islands

With the exception of the frog genus *Sooglossus*, which diverged from *Nasikabatrachus* in the K-Pg boundary with the separation between India and Seychelles, all datasets show that the diversification of extant clades in the Western Indian Ocean islands happened in the Cenozoic. This is no surprise as, due to the collision between India and Eurasia and the uplift of the Himalaya in the Neogene, the sea levels fell and created much less controversial land bridges between Africa and nearby continents. Madagascar's unequalled endemism also has

different origins for different taxa, but it seems to be almost as related to African taxa as to Australian and IAA taxa: for instance, whereas the pteropodid bats reached the Comoros, Madagascar, the Seychelles and the Mascarene islands through recent dispersal from India (Fig. 2.2d), the chameleons colonised Madagascar at least twice from Africa – and both dispersal events preceded further dispersal to Comoros, Eurasia and India. However, this contradicts suggestions made by deeper studies on the origin of Malagasy biota, which have pointed to a greater relationship to African clades (Yoder & Nowak 2006). Older studies also proposed such affinity (Simpson 1952; Krause 2003; Krause 2010), but here we also find a continuous invocation of land bridges to explain the colonisation of the island – even for the relatively short distances between Africa and Madagascar. Yoder and Nowak (2006) showed that most Malagasy taxa diverged recently, in the Cenozoic, which points to a very large number of dispersal events, even across the Indian Ocean. Thus, the uniqueness of the fauna of Madagascar could be due to such complex ancestry or the clades we sampled so far fail to represent the island's isolation. Boid snakes' presence in Madagascar has been discussed above and may represent a methodological failure of taxon sampling.

What our analyses consistently show is that Africa, Comoros, Madagascar, Seychelles and the Mascarene Islands are biogeographically connected: results for all taxa show both recent and old dispersal among these locations. Also, organisms show the same ease to travel from east to west as in the opposite direction through the Cenozoic; even though the existence of an east-west endemism boundary was neither evidenced nor denied by our results, we have shown that both extremities of the Indian Ocean have an intertwined history. Among the pteropodid bats, one particular dispersal to Africa gave rise to the Rousettinae-Epomophorinae-*Eidolon* diversification. This event is interesting because groups of *Eidolon helvum* Kerr, 1792 are able to migrate hundreds of kilometres (Fenton 2001) and it is possible that a group could have flown across the ocean with aid of the winds during the build-up of the MIOJet.

Further research should prioritise taxon inclusion, particularly plants and invertebrates. In order to capture more details on the paleobiogeography, the biogeographic history of more ancient taxa that are more diversified across the region should be attentively analysed. This study is yet in its beginning and there is still a long debate ahead before we can reach a consensus about the Indian Ocean's history. However, we are able to show that panbiogeographic scenarios such as the one proposed by Heads (2005) and Hocutt (1987) are too simple to accommodate reality. These results also show how poor our efforts on

conservation currently are: with such close relationship of such distant regions, the power of local actions to preserve recent diverged lineages is far from enough.

2.8 Conclusion

Some of the results we present here could be easily give rise to multiple interpretations. One of the main reasons for this was already discussed in the introduction: extinction is an important process that has been largely overlooked across biogeographic studies (Pillon & Buerki 2017). Take the result for the Boidae (Fig. 2.2a); event 2 represents the divergence between *Calabaria* and *Sanzinia-Acrantophis* and, according to the ancestral area reconstruction, the clade was originated in Africa and spread to Madagascar. However, given the distribution of the species in the phylogeny, a fossil taxon from Australia, India or Antarctica could easily change such inference from the dispersal Africa-Madagascar to the vicariance India-Madagascar. Such errors cannot be avoided but must be minded.

Given that these are preliminary results to a much more phylogenetically extensive analysis, the discussion will be carried out carefully to avoid false inferences. Currently, all analyses involve only vertebrates and over half of them are from squamates. Plants and invertebrates are likely the first to colonise a new ecosystem; analysing their taxa should give thus ideas about vicariant events o Gondwanan landmasses. Most of our results show a recent, Neogene history of diversification.

Chapter 3: The complex evolutionary history and phylogeography of *Caridina typus* (Crustacea: Decapoda): long-distance dispersal and cryptic allopatric speciation with comments on species delimitation

Most of this chapter has already been published as referenced in Bernardes et al. (2017)

Authors: Samuel C. Bernardes¹, Almir R. Pepato¹, Thomas von Rintelen², Kristina von Rintelen², Timothy J. Page^{3,4}, Hendrik Freitag⁵, Mark de Bruyn^{6,7}

¹Laboratório de Acarologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte 31270-901, Brazil. ²Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, Berlin 10115, Germany. ³Australian Rivers Institute, Griffith University, Nathan, Queensland 4111, Australia. ⁴Water Planning Ecology, Queensland Dept. of Science, Information Technology and Innovation, Dutton Park, Queensland 4102, Australia. ⁵Department of Biology, School of Science & Engineering, Ateneo de Manila University, Quezon City 1108, Philippines. ⁶School of Biological Sciences, Bangor University, Bangor LL572UW, UK. ⁷School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW 2052, Australia.

3.1 Abstract

Most of the evolutionary history of the freshwater shrimp genus *Caridina* is yet to be revealed in spite of all potential information contained in such diverse group and in their area of distribution, from Africa to Polynesia, known for its richness in biogeographical and geological history. Here we used nuclear and mitochondrial DNA to infer the phylogeographic and evolutionary history of *C. typus*, which has broadest distribution inside the genus and albeit it presents a high dispersive potential, questions have been raised on its phylogeographic structure and taxonomic status. We found three very distinct lineages that diverged in the Miocene. Molecular dating and ancestral range reconstruction are congruent with *C. typus*' early dispersal to Africa mediated by the Miocene Indian Ocean Equatorial Jet, followed by back dispersal to Australasia after the Jet's closure. Furthermore, different species delimitation methods agreed on treating each lineage as a distinct cryptic species. These events compose a complex evolutionary history in which ancient current systems and allopatric speciation

allowed secondary sympatry of cryptic species that may show unobserved ecological differences that provide reproductive isolation.

3.2 Introduction

The Atyidae (Crustacea: Decapoda: Caridea) are a family of small freshwater shrimps that occurs across all continents, except Antarctica. Fossil and molecular data show that the atyids are an old taxon, with estimates of divergence from other carideans ranging from the Permian to the Cretaceous (Porter et al. 2005; von Rintelen et al. 2012). The most speciose and most widely distributed atyid genus is *Caridina*, with over 300 species (de Grave & Fransen 2011) distributed from Africa to the Pacific islands, a range that is closely congruent with that of its type species, *Caridina typus* H. Milne-Edwards, 1837. This species has been shown to be highly tolerant to salinity, even in larval stages (Jalihal et al. 1994), and is never found too far from the sea (Johnson 1963; Suzuki et al. 1993), which could help to explain its wide distributional range. Small eggs and planktonic larvae (Jalihal et al. 1994) are also a feature that could contribute to *C. typus*' vagility (Hancock 1998). Despite this widespread distribution, *C. typus*, like other species of *Caridina* (Page et al. 2005; von Rintelen et al. 2007; Riek 1953), may actually comprise cryptic lineages. Phylogeographic and demographic studies on *C. typus* are limited, however Page et al. (2007) recovered strong support for a shallow intraspecific clade containing widely distributed *C. typus* specimens from Australia, Sri Lanka and several Pacific islands. A similar pattern of shared haplotypes was identified among populations of atyids isolated on distant islands of the Caribbean (Page et al. 2013). However, the opposite pattern has also been found for several recently identified (previously) cryptic species of *Caridina* (Woolschot et al. 1999; Chenoweth & Hughes 2003; Page et al. 2005), where deep intraspecific differences have been identified over very small scales. In fact, these cryptic species displayed different egg sizes, which could serve as an identifiable character (Page et al. 2005) and also as a rough indicator of dispersal capacity and resulting phylogeographic structure (Page & Hughes 2007), depending on how variable the trait actually is.

Caridina typus' distribution is interesting, as it comprises several areas with a very active geological history. First, the species is found in the Seychelles and in the Mascarene Islands, both part of the Mascarene Plateau: the former are granitic islands, fragments of Gondwana, whereas the latter are much younger volcanic islands (McDougall & Chamalaun 1969; Ashalatha et al. 1991). *C. typus* is also found through the Indo-Australian Archipelago (hereafter IAA), a region with a rich geological and biogeographical history. The IAA was

formed and shaped by the movement of the Australian Plate against the Sunda Shelf, and much of this region has been cyclically submerged and exposed through the Cenozoic, as Earth's ice-sheets have waxed and waned. Although thousands of islands arose from the Australian Plate's northward movement, many islands represented landmasses that were previously connected to Eurasia (reviewed by Lohman et al. 2011). Dispersal seems to be a very important process in this area, since most of the islands have never been connected to other land areas (van Oosterzee 1997), but vicariant events may also have been of importance with the opening of marine seaways, and orogeny that turned the landscape from a wide peninsula into the Archipelago we see today (Hall 2011; de Bruyn et al. 2013).

The history of the IAA has also influenced the biogeography of the Indian and Pacific Oceans. Before Sulawesi was formed, both oceans had a wide connection, but Goullan et al. (2008) found evidence to support the existence of the so-called 'Miocene Indian Ocean Equatorial Jet (hereafter MIOJet),' a strong westward current that would have restricted eastward transport. This current persisted until Sulawesi was completely established as a single landmass, an event that set in motion the currents we see today, and limited exchange between the two Oceans in the equatorial zone, via the Indonesian Throughflow.

Populations attributed morphologically to *Caridina typus*, given its broad distribution, provide an interesting model to understand the possible effects of this history on shaping the biogeography of the IAA, and its adjacent ocean basins. Given *C. typus*' planktonic larvae and the consequent ability to disperse across oceans, two alternative scenarios can be predicted on this species' genetic diversity: the first one is a widespread distribution with shallow geographic structure, showing that the species dispersive ability overcomes the geological history of its range; the second one is a species complex, where a deeper geographic structure reveals several cryptic lineages with more restricted distribution. The objective of this study, therefore, is to investigate *C. typus* phylogeography, species boundaries and the geographic distribution thereof, using multilocus molecular markers.

3.2 Materials and methods

3.2.1 Sampling

We sampled 117 *C. typus* specimens whose IDs and locations are listed in Supplementary Table S1 and plotted on the map in Figure 3.1. We sampled populations throughout the Indo-Australian Archipelago, including the Malay Peninsula, Borneo,

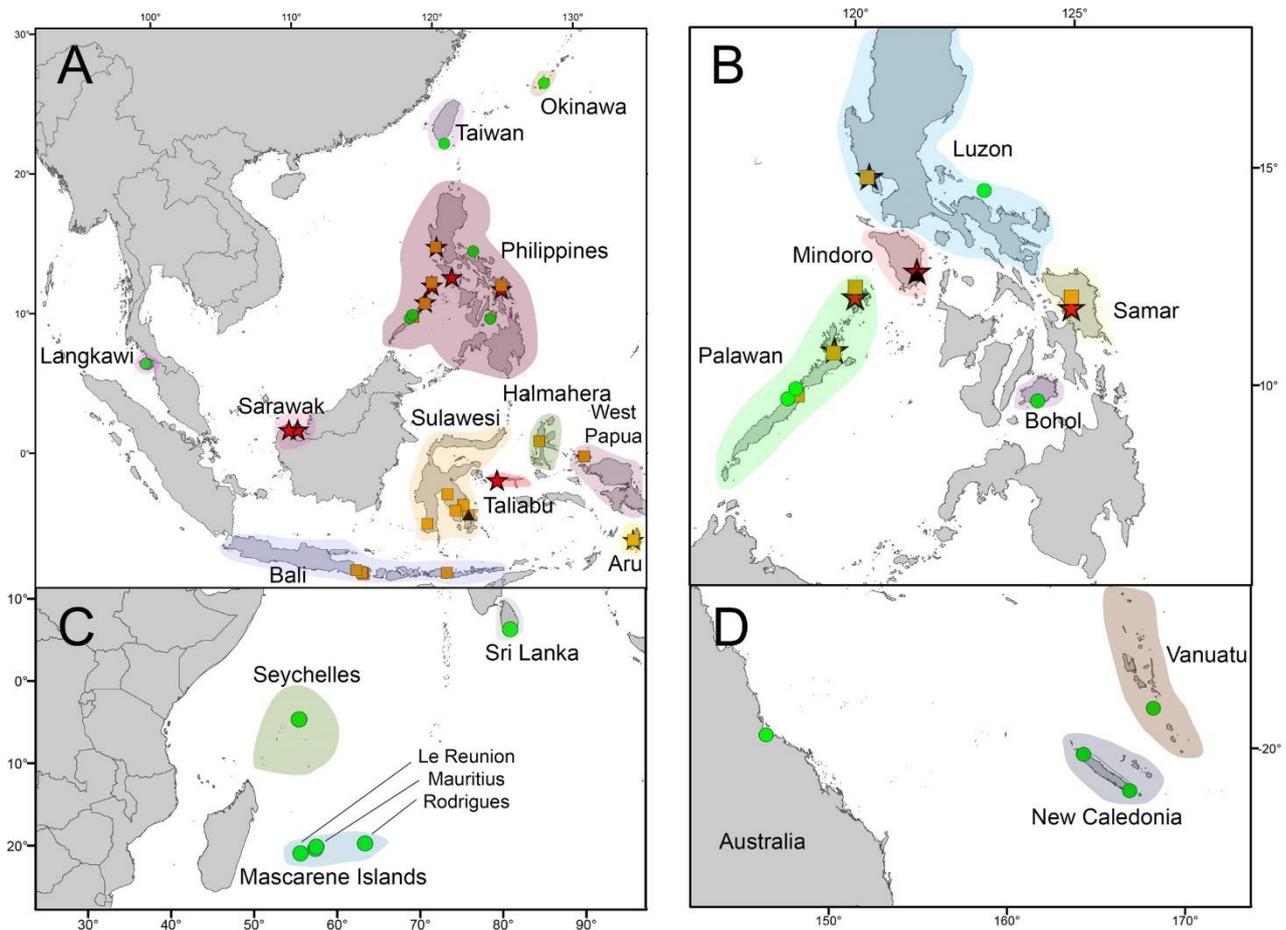


Figure 3.1: Sampling map of the Indo-Australian Archipelago (IAA) and Eastern Asia (a), with detail for Philippine islands (b); Western Indian Ocean (c) and South Pacific (d). The colours represent the regions and the symbols illustrate the three major clades (putative species): stars represent TAL, circles represent ARC and squares represent SUL. The triangles represent the other *Caridina* samples, *C. villadolida* and *C. opaensis*. Map and georeferencing were generated through ArcGIS Desktop: Release 10.

Philippines, Sulawesi, Java, Bali, Moluccas, Taliabu, Aru and West Papua, and also from outside the IAA, including Taiwan, the Seychelles, Reunion and Mauritius. The collections were conducted between 2003 and 2016, and samples were preserved in 100% ethanol. The samples are deposited in the Museum für Naturkunde Berlin (ZMB), Museum Victoria (CT_ArXX), Oxford University Museum of Natural History (OUMNH) and/or in the Molecular Ecology and Fisheries Genetics Laboratory in Bangor University. In our alignment, we added seven mitochondrial sequences from Genbank (see Supplementary Table S1) that expanded our sampling to Australia, Japan, New Caledonia, Vanuatu (these localities will be hereafter referred to as Pacific Islands) and Sri Lanka. In this study, we also included five

samples of *C. villadolidi* Blanco, 1939, which has historically been a synonym of *C. typus* var. *longirostris* De Man, 1892 (Chace 1997; Cai & Ng 2001) and two specimens of *C. opaensis* J. Roux, 1904, here used as outgroup.

2.4.2 DNA – extraction, amplification and sequencing

To extract DNA, Promega’s Wizard® Genomic DNA Purification Kit was used on a pair of pleopods following product instructions. We then amplified two mitochondrial loci, a 517 bp fragment of the subunit 16S of the mitochondrial rRNA and a 783 bp fragment of the subunit I of the cytochrome c oxidase (COI), and three nuclear loci, a 374 bp anonymous nuclear marker (ANM) named Ct33, a 330 bp ANM named Ct51, and a 368 bp fragment of the D3 segment of the subunit 28S of the nuclear rRNA. The standardised programs to amplify each marker are described in Table 3.1 as well as the primers used to amplify them. PCR products were purified with Beckman Coulter’s Agencourt AMPure® system except for CT33, which was purified with Affymetrix’s ExoSAP-IT® kit since CT33 PCR products, for some unknown reason, were frequently lost in purification through the AMPure system. The same primers used in amplification were used for sequencing. Sequencing was performed on an ABI 3130 sequencer using ABI Big Dye terminator chemistry (Applied Biosystems). All sequences were deposited in Genbank (accession numbers KY069328 - KY069787 and KY436221 - KY436224).

Table 3.1: Details on each locus amplification with primer references

Marker	Primer	Sequence	Program
16S	16S-F-Car ¹	5'-TGC CTG TTT ATC AAA AAC ATG TC-3'	95°C – 2 min; 40x [95°C – 30 s, 50°C – 30 s, 72°C – 30 s]; 72°C – 10 min
	16S-R-Car ¹	5'-AGA TAG AAA CCA ACC TGG CTC-3'	
COI	COI-F-Car ¹	5'-GCT GCT AAT TTT ATA TCT ACA G-3'	95°C – 2 min; 40x [95°C – 30 s, 45°C – 30 s, 72°C – 30 s]; 72°C – 10 min
	COI-R-Car ¹	5'-TGT GTA GGC ATC TGG GTA ATC-3'	
Ct33	CT33-F ²	5'-CCT TTC TAG ACG CAT CAA TGG-3'	95°C – 2 min; 10x [95°C – 35 s, 63°C – 35 s (-0.5°C/cycle), 72°C – 1 min]; 10x [95°C – 35 s, 58°C – 35 s, 72°C – 1 min]; 15x [95°C – 35 s, 52°C – 35 s, 72°C – 1 min] 72°C – 10 min ⁵
	CT33-R ²	5'-ATC TGA TTG GCT GGC TGA AT-3'	
Ct51	CT51-F ²	5'-GGG CTT TTA GCT AAG CTC TCG-3'	95°C – 4 min; 10 x [95°C – 30 s, 50°C – 45 s, 72°C – 2:30 min]; 18 x [94°C – 30 s, 45°C – 45 s, 72°C – 2:30 min]; 72°C – 10 min
	CT51-R ²	5'-GGC AGT TTC TTA TGG GCA TT-3'	
28S	28SD3AP ³	5'-CAA GTA CCG TGA GGG AAA GTT G-3'	95°C – 4 min; 10 x [95°C – 30 s, 50°C – 45 s, 72°C – 2:30 min]; 18 x [94°C – 30 s, 45°C – 45 s, 72°C – 2:30 min]; 72°C – 10 min
	D3-4283 R ⁴	5'-TAG TTC ACC ATC TTT CGG GTC-3'	

1- von Rintelen et al. 2007; 2- This study; 3- Reyda et al. 2003; 4- Lopes-Vaamonde et al. 2005; 5- de Bruyn et al. 2011.

3.2.3 Phylogenetic analyses

In order to assess *C. typus*' monophyly, multiple *Caridina* species' 16S sequences taken from Genbank were used in different combinations with our dataset. This analysis aimed only to give us an informative prior about the relationships within *C. typus* due to the inclusion of *C. villadolidi*, since our first limited trials depicted it as either a paraphyletic or a polyphyletic group. There was considerable difficulty amplifying all loci for some of the samples, thus we constructed several different datasets and many trials in order to obtain informative empirical priors, assess and minimise the effects of missing data, and maximise the information recovered from the sequences. From those initial trials, two datasets were retained for subsequent analyses: one comprising only mtDNA data (Dataset 1), and a second dataset (Dataset 2) encompassing all five alignments and all samples.

The ANMs and COI were aligned with ClustalW (Thompson et al. 1994; Larkin et al. 2007) as implemented in MEGA6 (Tamura et al. 2013). The mitochondrial ribosomal DNA fragment, 16S, was aligned through MAFFT (Kato & Standley 2013), while the nuclear ribosomal DNA fragment, 28S, was manually aligned through secondary structure (see Supplementary figure S2). The different alignment methods were due to the nature of the markers: autosomal and protein-coding markers were straightforward to align and homologous positions were easily identified by the position of nucleotides or codons. Ribosomal markers' homology is intertwined with the secondary ribosomal structure and more information from the marker can be accessed this way. Our 16S sequences were perhaps too short to exhibit the need for a manual alignment, but the presence of several gaps was problematic to ClustalW, thus we used MAFFT. Network 4.6 (Fluxus engineering) was then used to build a median-joining network (Bandelt et al. 1999) for each marker, taking gaps into consideration for all loci except COI that had sequences of different lengths, which were cut down to a similar size.

The best substitution model was chosen through the corrected Akaike Information Criterion (AICc) implemented in jModelTest 2.1 (Darriba et al. 2012). In order to choose the best molecular clock model, we ran multiple analyses for each marker separately with either a strict or relaxed molecular clock (both exponential and log normal models were tested), and we then used TRACER v1.6 (Drummond et al. 2012) to compare log values using an AICM approach as described by Baele et al. (2012).

3.2.4 Analyses with Datasets: priors and model setting

For the monophyly assessment trials, we ran a coalescent clock based analysis in standard BEAST 2 v. 2.3.0 (Bouckaert et al. 2014) using a reversible-jump method to choose the best substitution model (Huelsenbeck et al. 2004; Bouckaert et al. 2013). In addition to the clock based analysis, we also conducted a non-clock analysis with MrBayes 3.2.5 (Ronquist & Huelsenbeck 2003) to assess the rooting recovered by BEAST2. The input file was run for 10 000 000 generations, with four simultaneous chains, and sampling every 1 000 generations.

For Datasets 1 and 2, we also ran a coalescent clock based analysis for these samples, but this time we used *BEAST (Bouckaert et al. 2014) in order to use multiple alignments. A reversible-jump method was used to choose the best substitution model and the runs were visually diagnosed through TRACER. The molecular clock models used for each locus were selected through AICM analysis: relaxed lognormal clock for the mitochondrial markers and strict clock for the nuclear markers. A constraint in the ingroup monophyly (*C. typus* specimens) was also used based on the results of the monophyly assessment and only proper prior distributions were chosen, with the exception of the clock rate prior for strict clocks, which were set to a proper bounded uniform distribution. Each analysis was run for 500 000 000 generations and sampled every 10 000 generations. Runs were repeated 3-5 times to assess the repeatability of the results, and a 50% burn-in was used in *TreeAnnotator*.

A substitution rate of 1.1-1.3% per million years for COI was used to date the cladogenesis events. This rate was taken from the divergence rates presented by Knowlton et al. (1993) for a Caridean shrimp based on the rise of the Isthmus of Panama and it was previously used for a *Caridina* species by Hurwood & Hughes (2001). Since it is somewhat close to the COI rate pointed by Brower (1994) it should be a fairly conservative approach to estimate the node ages. In the preliminary analyses and trials, several inconsistencies between the dates of the COI trees and the multilocus trees were found, as well as difficulty in MCMC convergence for the multilocus analyses. In order to avoid age inflation or deflation and to facilitate the convergence of the runs, informative priors on the relative rates of the other loci were estimated in Garli (Zwickl 2006) following the method described in Marshall et al. (2016).

The priors and models were tested in several trials with different combinations of priors' distributions, and markers. All analyses in BEAST, MrBayes, jModelTest and Garli were run in CIPRES (Miller et al. 2010).

3.2.5 Species delimitation analyses

To use the Generalised Mixed Yule Coalescent (GMYC) approach on species delimitation (Fujisawa & Barraclough 2013; Pons et al. 2006), we needed a mtDNA alignment that included only unique haplotypes, since this approach was developed for single locus use and is sensitive to the length of the branches. The Bayesian approach of the same method (bGMYC) introduced by Reid & Carstens (Reid & Carstens 2012) was also used. Even though both GMYC and bGMYC are different mathematical implementations of the same model, and the latter is known for reducing uncertainty of the former, we preferred to include both sets of results for our data in order to analyse and identify possible causes for differences between them. For the same reason, trials were made for separate 16S and COI trees to check for congruence between loci.

For Dataset 2, we used the package STACEY (Jones et al. 2014; Jones 2016) implemented in BEAST2 to infer minimal clusters from each dataset. We used very conservative priors and favoured a single cluster (e.g. high values for collapse weight (ω) and smaller values for collapse height (ε)) in order to avoid overestimating the number of species. The analysis was repeated twice, and run for one billion generations sampling every 100 000 generations.

Another method used was the discovery method implemented in BPP v.3.2 (Yang & Rannala 2010; Yang 2015). The results from all the species discovery analyses were used as input to the validation method of BPP. Both BPP approaches had MCMC samples taken every 2 iterations with a total 2 000 000 samples and a 10 000 iterations burn-in plus a *thetaprior* = 2 2000 and *tauprior* = 2 1000.

3.2.6 Demographic and phylogeographic analyses

Geneland (Guillot et al. 2005, 2012) was used as an approach to delimit populations inside each main clade. Haplotype networks were inconclusive and we did not want to apply geographic distance. Due to sympatry, each clade was run separately and due to the amount of missing data present for nuclear markers, only mtDNA data were run. Three independent runs were set with 10 000 000 iterations with a maximum of 20 populations. We then used Arlequin 3.5 (Excoffier & Lischer 2010) to perform an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) for both mitochondrial and nuclear loci. The same software was used to conduct two neutrality tests (Tajima's D and Fu's F_S) in order to identify any signatures of

demographic change. Again, neutrality tests were conducted only for mitochondrial loci due to excessive amounts of missing data for nuclear markers. Extended Bayesian skyline plots (EBSs) (Heled & Drummond 2008) were also used to assess the demographic history of each clade, run in BEAST2 with the same respective parameters for Dataset 1 and 2 phylogenetic analyses.

To reconstruct the ancestral range, we used RASP (Yu et al. 2015) to apply the BayArea model. This model was chosen due to the feature of implementing distance to the probability of dispersal given that our studied organism is quite widespread. BayArea was run for 5 000 000 generations sampling every 1 000 generations with a TRUE setting for geographic distance power. For reasons of convergence and limitations in the number of locations accepted by BayArea, the populations identified by Geneland were taken into account for each sample's location assignment. For instance, the samples from Sri Lanka were included with Langkawi. We also used the Bayesian discrete phylogeography method (Lemey et al. 2009) implemented in BEAST with the same parameters as applied to our phylogenetic analyses. Convergence and number of locations was not a problem with this analysis, but to ease comparability between methods, we used the same location assignments as in BayArea. The phylogeographic analyses were run only for the combined multilocus Dataset 2.

3.3 Results

3.3.1 DNA – extraction, amplification and sequencing

While the mtDNA markers were fairly easy to amplify and sequence, the nuclear markers (with the exception of Ct33) were very difficult to amplify. The missing data for ANM Ct51, we believe, might have been caused by the absence of a primer binding site, since the samples that were not sequenced were either from the same location (Sarawak), or were shown as belonging to the same clade in the phylogenetic reconstructions. The partial ribosomal sequence for 28S has the largest amount of missing data due to inability to amplify the locus for many samples. The transportation of the bulk of the samples from the UK to Brazil was delayed by customs issues and may have led to sample degradation.

3.3.2 Phylogenetic analyses

The analyses of many *Caridina* species revealed *C. typus* to be a paraphyletic group (Supplementary Figure S3), consisting of three morphologically similar clades of “*C. typus*”

and the morphologically distinct *C. villadolidi*. Since morphological data suggest *C. villadolidi* is a separate species, and as it was only sequenced for 16S (see Supplementary Table S1), we decided to concentrate on clades of *C. typus* morphotypes. Therefore, we constrained the monophyly of *C. typus* to these three clades and excluded *C. villadolidi*, which also had the effect of improving chain convergence. We will revisit *C. villadolidi* once the genetic and morphological relationships of *C. typus* become clearer. Figure 3.2 shows the final phylogenetic tree obtained for this data set. Both Dataset 1 and Dataset 2 recovered the same clades with some minor differences in the placement of a few individuals. Three main clades were identified and named according to their range: ARC (samples from Mauritius, Seychelles, Sri Lanka, Australia, Pacific, Langkawi and Philippines, comprising more or less an arc that surrounds the IAA), SUL (samples from across the IAA, centred on Sulawesi) and TAL (samples from Philippines, Malaysia, Taiwan and Taliabu). The dates obtained from Dataset 2 were much more recent compared to those obtained from Dataset 1 (both are represented in Fig. 3.2). Trees from Dataset 1 (mtDNA data) exhibited longer internal branches at the origin of each clade, with shallow differentiation within the clade. Internal branch lengths were much shorter for Dataset 2 trees.

Figure 3.3 shows the haplotype network for the mitochondrial 16S. Haplotype networks for all five markers can be seen in Supplementary Figure S4. The ANM Ct51 (Supplementary Figure S4d) dataset did not include a sample from the ARC clade. The mtDNA networks showed high distinction between clades and, as expected, COI (Supplementary Figure S4b) showed a considerable degree of population structure. The 16S network (Fig. 3.3) was more conservative and exhibited very limited structure in SUL. However, it showed a structured pattern in ARC that was congruent with the high differentiation found in COI. The ANMs Ct33 and Ct51 (Supplementary Figure S4c and d) had the most conserved pattern and exhibited a star-like network and no structure.

3.3.3 Species delimitation analyses

All the different species delimitation approaches, either single or multilocus, achieved the same result, with the three main clades identified as different species (see discussion). Even

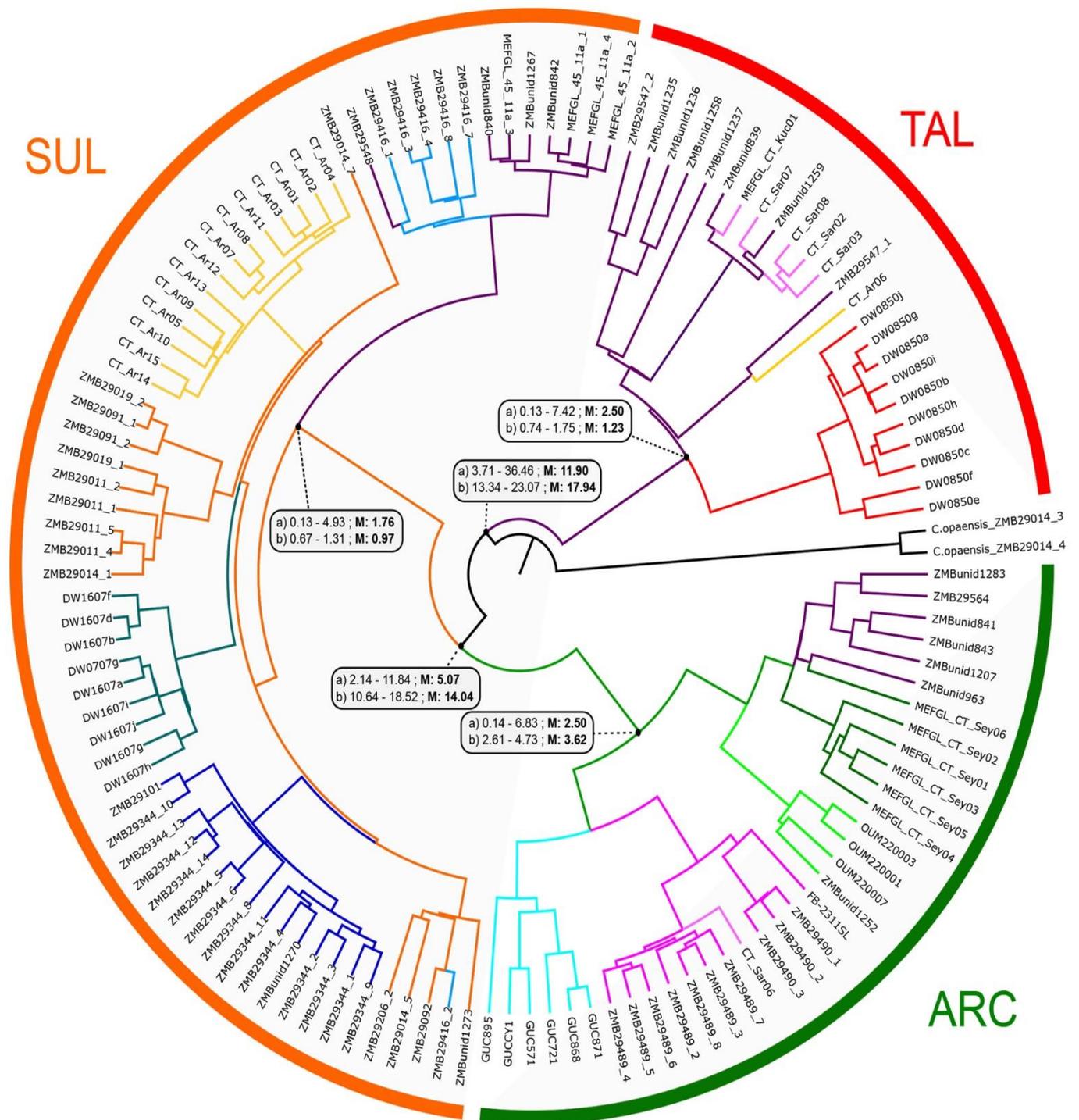


Figure 3.2: Time-calibrated phylogenetic and phylogeographic relationships in *C. typus* rooted using *C. opaensis*. Dates are represented as millions of years before present. The main nodes, i.e., those that represent divergence between the major clades and origins of the diversification within them, are labelled with 95% confidence intervals for their dates as well as its median (M). Each label carries the date obtained through data from all markers (a) as well as the one obtained solely through mtDNA data (b).

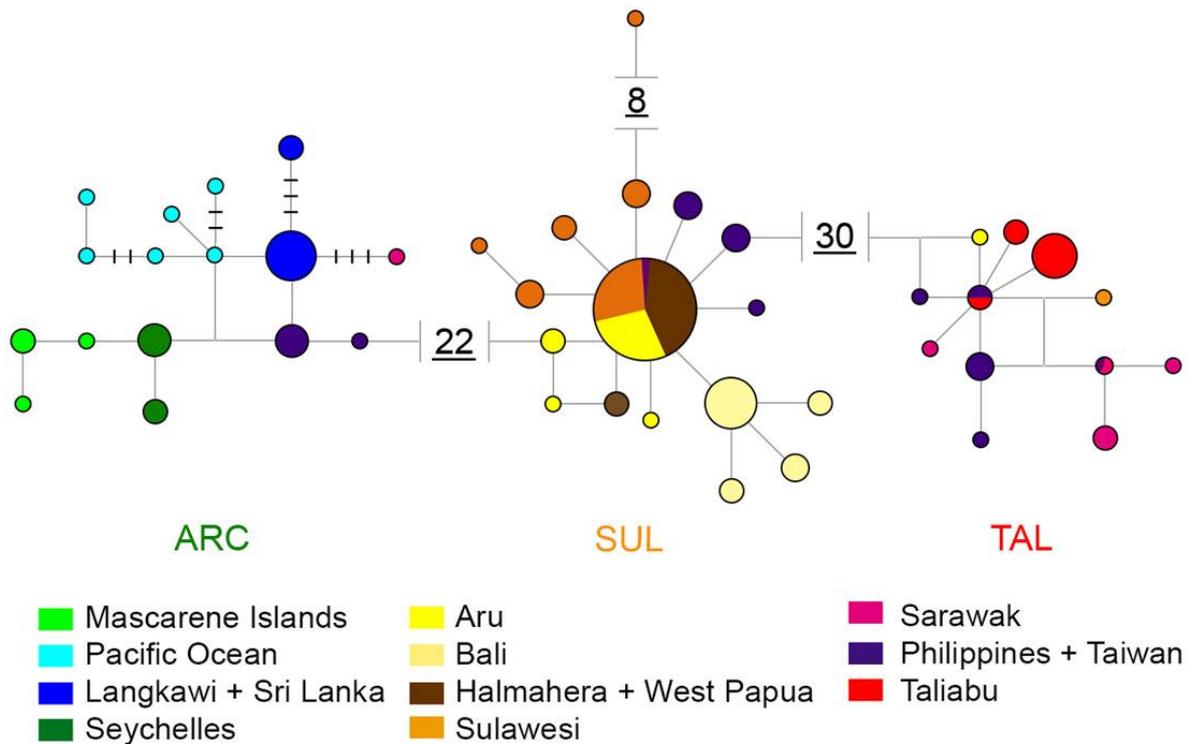


Figure 3.3: Mitochondrial 16S haplotype network. A single substitution step is represented by a single line connecting two haplotypes. Any number of steps between two and five is represented by the number of dashes on the line connecting two haplotypes. Any distance higher than five steps is represented by the written number of steps. Non-sampled intermediary haplotypes are represented by empty vertices connecting three or more haplotypes. Each colour represents a region according to the legend.

GMYC, that was conducted with a fairly different tree and is known to be prone to overestimation of the number of species due to difficulty of the model to assess population structure versus species boundaries (Satler et al. 2013; Carstens et al. 2013), recovered these three distinct clades in all trials with 16S and total mtDNA data (95% confidence interval = 3-4). Trials with only COI trees had TAL, SUL and ARC divided in two, five and four species respectively (total species number = 11; 95% confidence interval = 4-13). These trials with COI had much higher values for the likelihood ratio test ($LRT = 1.96 \times 10^{-4}$) than the ones that involved 16S ($LRT = 6.94 \times 10^{-9}$) and 16S+COI ($LRT = 5.21 \times 10^{-10}$).

BPP showed differences in the results for its discovery and validation approaches. In the validation approach, a very high probability for the existence of four different species (three *C. typus* species plus *C. opaensis*) was found (*posterior probability* (pp) = 0.9722) with a high support for every node ($pp = 0.9723$ for the separation of TAL and 0.9722 for the divergence between ARC and SUL). However, even though the same number of species was obtained by

the discovery approach of BPP, the probability was much lower ($pp = 0.68601$), a value that corresponds to the probability of ARC and SUL being different species. The reduced value was due to the relatively high probability of ARC and SUL belonging to the same species ($pp = 0.30195$) compared to the probability of TAL being a separate species ($pp = 0.98796$).

The first trials with STACEY showed a noticeable trend toward more speciose results (10-17 species) but they were far from convergence with very low ESS values. As stated before, we decided to use very conservative priors and this also appears to have enabled convergence. The final result revealed three main clusters besides the outgroup (see Supplementary Figure S5).

3.3.4 Demographic and phylogeographic analyses

Geneland analyses were somewhat problematic due to secondary sympatry – especially in the Philippines – and due to the nuclear markers: when run as a unique set, very different samples from the same place would be placed in the same population and no convergence would be obtained. The SUL clade was particularly troublesome due to the number of repeated haplotypes across the IAA islands. Running each clade separately, 12 populations were found across the three clades: SUL: Sulawesi, Bali, West Papua and Halmahera, Aru and Philippines_SUL; ARC: Africa (samples from Seychelles and Mascarene Islands), Langkawi (which includes the sample from Sri Lanka), Pacific Ocean (samples from Australia, Japan, Vanuatu and New Caledonia) and Philippines_ARC (which includes Taiwan); and TAL: Taliabu, Sarawak, and Philippines_TAL.

Table 3.2: Results of a two-level AMOVA of genetic differences for both mtDNA and nDNA sequences of the three main clades of *C. typus* species-complex

Group	mtDNA		nDNA		p-value	
	Among populations (%)	Within populations (%)	Among populations (%)	Within populations (%)	mtDNA	nDNA
ARC	66.51	33.49	88.96	11.04	$p < 0.00001$	0.00196
SUL	20.71	79.29	6.44	93.56	0.00391	$p < 0.00001$
TAL	99.37	0.63	49.70	50.30	$p < 0.00001$	$p < 0.00001$

Notes: Entries represent the percentage of the total variance that is explained by variation within and among the populations defined by Geneland

AMOVA results are described in Table 3.2. All three clades exhibited consistent signs of geographic structure in their populations, even though there were differences in the partitioning of genetic diversity between them: whereas ARC and TAL had greater variation

Group	Population	Tajima's D (p-value)	Fu's F_S (p-value)
ARC	Africa (Seychelles-Mascarene)	0.3845 (0.6770)	-1.1814 (0.1270)
	Langkawi	-0.9762 (0.1510)	3.3129 (0.9380)
	Pacific Ocean	-0.1247 (0.4780)	-3.2830 (0.0150)*
SUL	Philippines_ARC	-1.1320 (0.1630)	-0.8584 (0.0810)
	Aru	-0.7985 (0.2260)	-1.2286 (0.2690)
	Bali	-1.7625 (0.0270)	-2.8465 (0.8780)
	Halmahera-West Papua	0.4964 (0.7350)	-1.6241 (0.2040)
TAL	Philippines_SUL	-1.4208 (0.0700)	6.3649 (0.9830)
	Sulawesi	-1.8258 (0.0110)	16.8897 (1.0000)
	Philippines_TAL	-0.5607 (0.3200)	-2.5803 (0.0120)*
	Sarawak	-0.9726 (0.0980)	-0.8292 (0.0970)
	Taliabu	-0.0203 (0.4920)	-0.7006 (0.1750)

Notes: Values in bold represent results where $p < 0.05$. Values marked with * represent Fu's F_S results where $p < 0.02$

structured among populations, the SUL clade showed higher variation structured within populations. All results described by AMOVA were highly significant.

Neutrality tests (Table 3.3) showed conflicting results between Tajima's D and Fu's F_S . Tajima's D values were significantly negative and indicated departures from neutrality for the Pacific Ocean, Bali and Sulawesi, but only the Philippine_TAL population showed significant values for neutrality departure for Fu's F_S under a significance level of $p < 0.05$. Nonetheless, Fu (Fu 1997) showed that an appropriate level of significance for F_S is $p < 0.02$, which would make nearly all F_S values non-significant. Samples from Bali show clear signs of recent expansion in the haplotype networks for both mitochondrial markers (Fig. 3.3 and Supplementary Figure S4b), but neither the Sulawesi nor the Pacific Ocean populations exhibited the same pattern. The latter has in addition restricted sampling, with only a few individuals sampled from distant islands.

Figure 4 shows the EBSPs for both Datasets 1 and 2. Both ARC and TAL clades show a slight and continuous recent growth for mtDNA (Fig. 3.4a and c), whereas SUL shows an

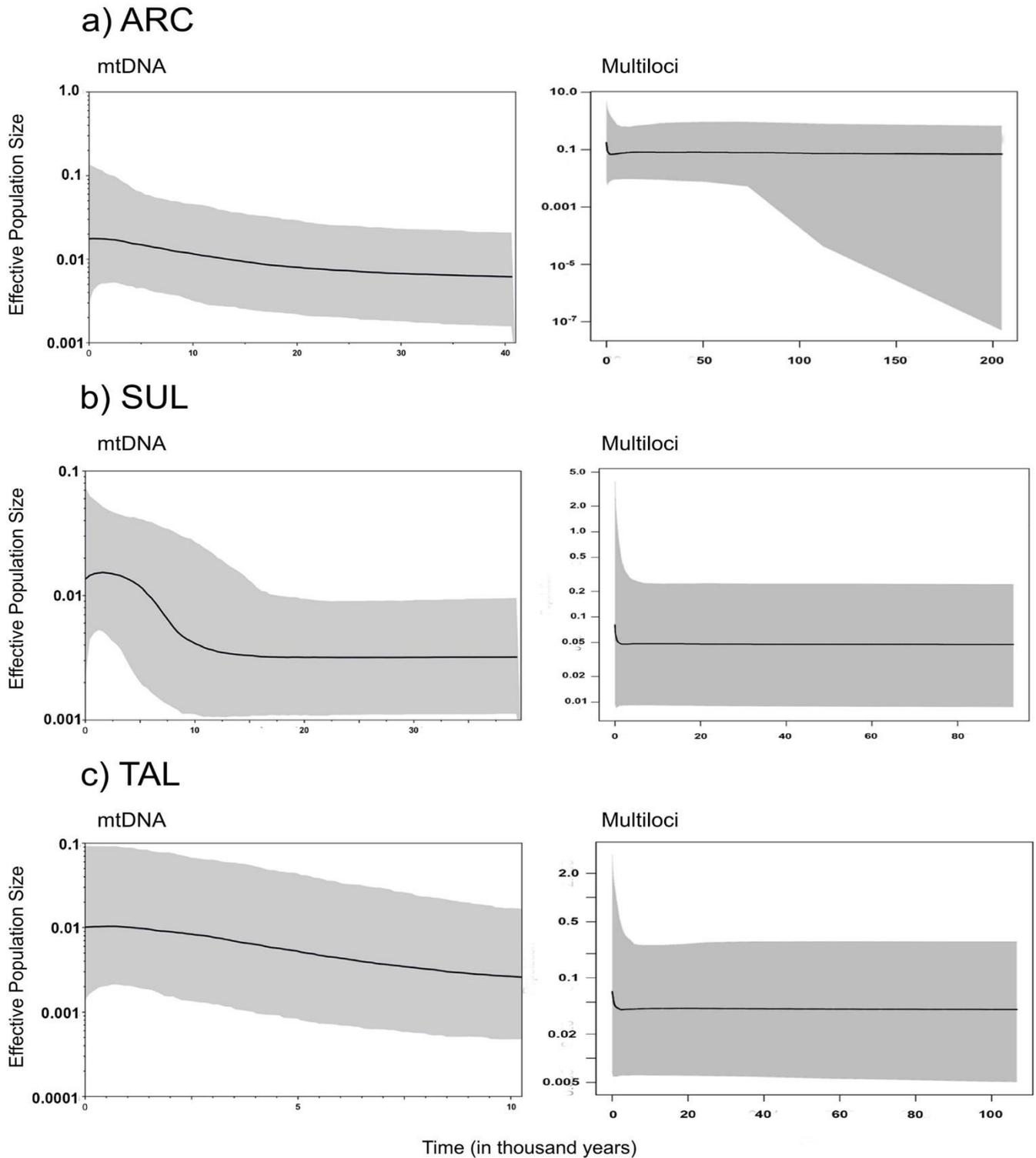


Figure 3.4: Extended Bayesian skyline plot for mtDNA (left) and multiloci (right) data for each main clade of *C. tybus*: ARC (a), SUL (b) and TAL (c). The bold line indicates the median whereas the grey area represents the 95% confidence interval. The dates were calibrated with COI substitution rate and are shown in thousands of years before present.

abrupt recent expansion (Fig. 3.4b). When all data is included, all three clades seem to be stable over the last 100 000 years (at least). However, the 95% confidence interval for ARC may suggest that either our markers or our sampling was not able to capture a strong expansion that is older than 100 000 years.

The result of BayArea (Supplementary Figure S6) shows high uncertainty for deeper nodes and the result shown for *C. typus*' ancestor was probably caused by the fact that the outgroup is from Sulawesi and that all three clades are present in the Philippines. Yet, the more recent nodes show enlightening patterns. First, the node for the ancestor of TAL (node 248 in Supplementary Figure S6) has Philippines as the most probable range – and it has a high posterior probability considering that all proposed ranges have Philippines included in them (Philippines-Taliabu: $pp = 0.527$; Philippines-Taliabu-Sulawesi: $pp = 0.0879$; Philippines: $pp = 0.0852$). Taliabu also shows high probability of being part of the ancestral range. Using the same principle, Africa is the most probable ancestral range for ARC's ancestor (node 159 in Supplementary Figure S6) and, though most of the possible ranges have two areas (Africa: $pp = 0.1944$; Africa-Philippines: $pp = 0.1278$; Africa-Pacific: $pp = 0.1092$; Africa-Langkawi and Africa-Sulawesi: $pp = 0.0639$), Africa consistently appears as part of the ancestral range in different combinations. Finally, node 224 (SUL's ancestor) shows Halmahera as a consistent part of the proposed ancestral ranges, but, similarly for Taliabu in node 159, Sulawesi also has high probability of comprising the ancestral range. All the probabilities, shown as pie charts, are displayed in Supplementary Figure S6.

The results of BEAST's Bayesian discrete phylogeography were inconclusive. The discrete phylogeographic analysis indicated that the first clade to diverge had a similar distribution to the ancestral range and was highly affected by sympatry: the most probable ancestral range of a node was the one that was shared by its branches. Moreover, some of the most external branches (e.g. the most recent common ancestor of the samples from the Pacific) included Taliabu as a possible part of the ancestral range, even though all Taliabu samples pertained to TAL. That might be because the tree's dates are too ancient for the method to find a "measurable footprint" (Lemey et al. 2009). The results can be seen in Supplementary figure S7.

3.4 Discussion

Our analyses revealed three major clades that cannot be readily explained by the IAA's geological history alone. As expected for such a vagile species, even though each clade exhibits a fair degree of population structure, the populations are not restricted to a single landmass, and often comprise samples from several islands (Fig. 3.1 and 3.3). AMOVA tests (Table 3.2) show that the clade SUL, which comprises most of the IAA, exhibits greater differentiation within populations than between populations, which is well illustrated by the haplotype network recovered for COI (Supplementary Fig. S3b).

This is in agreement with the findings of Fujita et al. (2016), whose conclusions were based on less conserved markers and revealed a mixture of two distinct lineages across Japanese islands. Even though the COI region they sequenced overlaps ours by only 167 bp, both of their lineages are placed inside ARC, as we would expect given the similar placement of our Okinawa specimen (see Supplementary Fig. S7).

Thus, based on our data, the three main clades, ARC, SUL and TAL represent independent lineages each with its own history of diversification, expansion and differentiation. Results therefore reflect, at different scales, elements of both of our predictions: a species complex, where a deeper geographic structure reflects restricted distributions, while at a finer scale each of these lineages display shallow phylogeographic structure, suggesting dispersive ability overcomes geological history within each lineage's range.

TAL shows a high degree of divergence to the other clades, being distant enough to be consistently recovered as a separate clade in the nDNA networks (Supplementary Fig. S3). Moreover, TAL may not be reciprocally monophyletic relative to the other two clades, as it appears to be more closely related to *Caridina villadolidi* (but note, weakly supported in Supplementary Fig. S2), which would make *C. typus* a paraphyletic taxon. The EBSP for TAL (Fig. 3.4) shows a more or less stable population. The slight growth trend recovered by mtDNA could represent a recent expansion in the Philippines, as these samples showed a significant negative value for Fu's F_S . BayArea indicated an ancestral range centred in the Philippines for TAL, with the addition of Taliabu. These results may be an artefact of sampling since these methods can only assess the ancestral ranges if they are included in the current range and sampled – and, in this case, these areas are also time dependent as the Philippine islands were far south of their current position as recently as 5 Ma (Hall 1996; Lohman et al. 2011). Alternatively, as there is clearly a north-south structure (Fig. 3.2), one possibility would be a Cenozoic range located to the south or east of the Philippines current position (Fig. 3.5a), a

scenario that agrees with Renema et al.'s (2008) "Miocene Australasian hotspot of biodiversity".

The lower divergence between ARC and SUL for the nuclear markers may be due to limited sampling for the ARC clade: samples from continental Africa or Asia were not available for analysis, nor have we sampled exhaustively in the Pacific, and both regions could host genetically distinct ARC populations. It is important to note that no ARC samples would amplify at Ct51, which could be a sign of consistent null alleles across the clade and/or mutations in priming sites, and thus it could be speculated that the degree of differentiation between ARC and SUL for that marker should be equal to or higher than the pattern seen in 28S (Supplementary Fig. S3e). Even though the nuclear markers recovered ARC and SUL as different clades (Fig. 3.2), node dates differed from those recovered solely from mitochondrial data. While mitochondrial data suggest an ARC-SUL divergence around 14 Ma, multiloci analyses indicated a much more recent cladogenesis at ~ 3.8 Ma. It must be stressed that molecular dating is very imprecise with very large confidence intervals, and is presented to give a general idea of the geological age of divergences and to provide discussion points for biogeographic hypotheses. These dates, in spite of their difference, however, provide a working hypothesis associated with the MIOJet. According to Goullan et al. (2008), the MIOJet began ~ 14 Ma and lasted until ~ 3.5 Ma. Once the MIOJet was established, individuals from an ancestral population (most likely SUL or TAL) could have dispersed westwards (Fig. 3.5c) and colonised the Seychelles, and later the Mascarene Islands, whose oldest island, Mauritius, began to emerge around 8 Ma (McDougall & Chamalaun 1969). At ~ 3.5 Ma, with the closure of the Indonesian Gateway and the termination of the MIOJet, contact between ARC and SUL would have been terminated.

With the closure of the MIOJet, ARC individuals may have become able to move eastwards and to colonise Malaysia and the Pacific (Fig. 3.5d). Results from BayArea analyses support that scenario, with Africa a likely part of ARC's ancestral range. The Philippines is interesting in this scenario: BayArea shows a considerable probability of the Philippines being part of ARC's ancestral range, yet in spite of the amount of missing data, a clear separation was evident between ARC samples from the Indian and Pacific Oceans (Fig. 3.2). The Philippines could indeed have been part of the ancestral range as the starting point for ARC's MIOJet-mediated dispersal (Fig. 3.5b), but if this were the case, ARC-SUL divergence would have occurred before ARC's dispersal throughout the Indian Ocean.

EBSPs for mitochondrial data (Fig. 3.4) depict an abrupt expansion for SUL and, although not statistically significant, a slight steady growth trend for both ARC and TAL, whilst the plots obtained for all markers show signs of growth only for ARC. This is congruent with the MIOJet scenario, where ARC should have had a somewhat constant spatial expansion for the past 14 Ma, and more so over the last ~ 3.5 Ma. Fu's F_S shows a significant sign of recent expansion for the samples from the Pacific Islands (Australia, Japan, Vanuatu, New Caledonia; Table 3.3), which is also in agreement with the MIOJet scenario, but could be artefactual due to limited sampling for that region. The same significance is not found for Tajima's D value, but F_S should be more sensitive to departures from neutrality (Fu 1997).

The mtDNA EBSP for SUL (Fig. 3.4) exhibits a signature of a recent expansion that agrees with patterns found using AMOVA and haplotype networks. Nonetheless, the lack of structure and high gene flow between islands may be more in agreement with constant demographic size depicted by the total evidence from all five markers. Departure from neutrality in Tajima's D was only significant for the Sulawesi population, but it does not show significant F_S values. These results could represent the colonisation of several Indonesian islands following the end of the MIOJet, since the dates found for the origin of SUL are very recent, but could also be related to female philopatry.

The question raised by these patterns and findings is, are there any ecological or developmental differences between each clade? Atyids' life history could be associated with their egg size (Jalihal et al. 1994; Hancock 1998), and in this context, *C. typus*' hardy planktonic larvae can be related to its dispersal ability. However, while ARC is exceptionally widespread across both the eastern and western hemispheres, SUL and TAL are restricted to the IAA, which could raise questions on the differences between each clade's eggs and larvae. Differences of that kind have been shown in other studies on *Caridina* (Page et al. 2005), even though differences in vagility were not as obvious in that case. Fujita et al. (2016) showed that Japanese *C. typus* has a considerable dispersal ability, but, as far as we can tell, their samples are restricted to our ARC clade. In the Early Miocene, when the Indonesian Gateway was wide open, eggs and larvae could have been carried toward the rest of the IAA from *C. typus*' original range (Figure 3.5a) (see Lohman et al. 2011) and established the ancestral ARC and SUL populations (Fig. 3.5b), probably in the southerly part of those ranges (given the BayArea results), with Halmahera and Sulawesi the most likely ancestral ranges. The dates for TAL's divergence from ARC-SUL range from ~ 18 Ma (mtDNA) to ~ 4 Ma (based on the multilocus dataset), which may indicate that gene flow between TAL and the ARC-SUL's ancestral

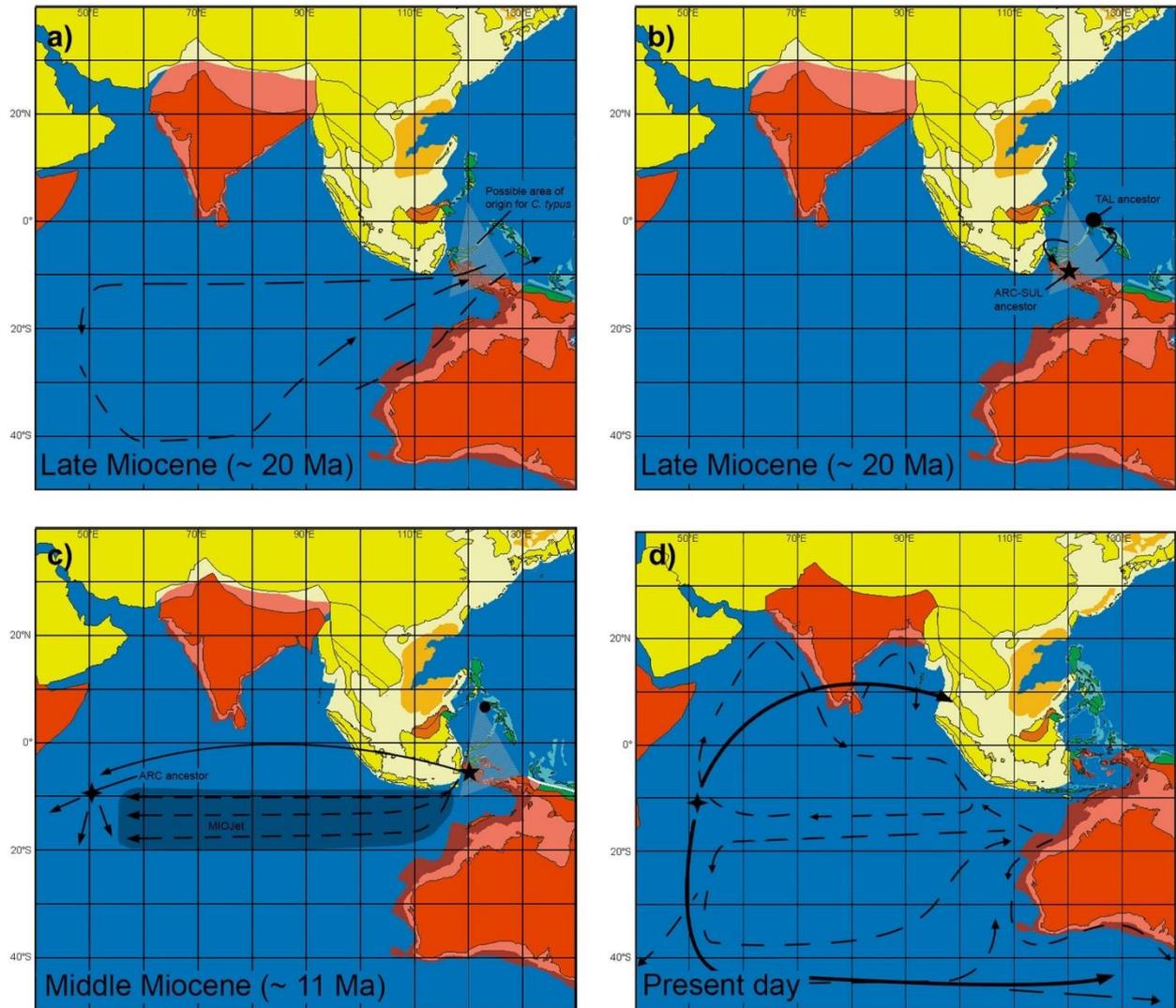


Figure 3.5: An illustrated suggestion for the biogeographic history of *C. typus*. In (a), the grey dashed triangle represents the possible ancestral location for the population that would originate based on TAL's ancestral range inference. The ancestral population would have spread to other locations in the precursor IAA (b) probably westwards, to Taliabu, the precursor islands of Sulawesi and/or, based on ARC's ancestral range inference, the Philippines. With the establishment of the MIOJet, individuals from this ARC-SUL ancestor would have spread to the Indian Ocean and initiated the population that would originate ARC (c). The closure of the MIOJet would finally establish currents as they are today and permit the colonisation of Langkawi and the Pacific by ARC. The dashed arrows represent surface currents based on Gourelan *et al.* (2007) and bold arrows represent postulated dispersal events. The dots represent possible points of establishment and isolation of TAL; the star indicates possible points of origin of the ARC-SUL clade; the compass rose indicates the most probable point of origin for the ARC clade. Coastal shelf areas are marked in a lighter tone than the continental areas. Paleomaps were modified from Hall (2012).

population became restricted by the MIOJet and by the formation of Sulawesi and other islands of the IAA, as the SE Asian Gateway narrowed significantly by 5 Ma. If the same dispersive ability was present in all three clades, ecological exclusion could explain why ARC was not able to recolonise the IAA (except for the Philippines), and why SUL and TAL's individuals did not also establish populations of their own outside the IAA.

Sukumaran and Knowles (2017) have pointed out that coalescent-based methods of species delimitation may not track speciation, but rather tracks emergent lineages that represent population structure, thus overinflating the number of species identified. Even though their simulations were run with BPP – which was a method used in this paper – and albeit we agree that species delimitation should be approached with caution, two (GMYC and STACEY) of the three methods used here were based on mixed models. Due to the fact that no morphological differences between the three clades described here have been identified to date (pers. comms. Werner Klotz & Sammy De Grave), we were very conservative in the analysis and favoured a single species where possible. Methods such as GMYC and BPP depend solely on the input tree and sequences, but STACEY depends on convergence, and thus on priors. While setting informative empirical priors, we conducted several trials favouring different numbers of species, but we aimed to design a robust test for the null hypothesis that *C. typus* comprises a single monophyletic species. The three *C. typus* clades identified in our analyses represent very distinct lineages and the fact that all species delimitation approaches, even those that have been shown to be prone to overestimation like GMYC, agreed on the final result strongly suggests the samples included here represent three distinct species. We do, however, acknowledge the risk of a false positive (Sukumaran & Knowles 2017) from the methods used here, and highlight the need for future morphological and ecological studies to test these results.

Many levels of sympatry were found: two individuals from Aru and from Sarawak were consistently recovered as belonging to TAL and ARC respectively, and individuals from the Philippines were roughly equally divided between the three clades. This study did not find any conclusive evidence for hybridization between clades, but the majority of the Philippines' samples were sequenced only for mitochondrial 16S, which can not reveal hybridisation. Efforts toward a more detailed morphological description of the specimens could reveal for the Philippines, for instance, a similar scenario to that of von Rintelen et al.'s study on Sulawesi lakes (von Rintelen et al. 2010), that found similar populations diverging along the same habitat, which in this case could occur following secondary contact. Alternatively, we may find a late stage in radiation (Streelman & Danley 2003) like that found for *C. ensifera* (von Rintelen

et al. 2007), with prezygotic isolating mechanisms playing a role. Finally, we agree with Page et al.'s (2005) suggestion that egg and larva size assessment between clades should be conducted, and may inform on differences in vagility, and thus perhaps evolutionary patterns, among the three clades.

3.5 Conclusion

This study has shown that *C. typus*, as currently defined under morphological criteria, comprises three distinct clades with unique evolutionary histories. One clade, ARC, may be a useful model taxon for future research due to its extremely wide distribution and clear separation between populations from the Indian Ocean and those from the Pacific Ocean. The MIOJet may have played an important role in this clade's history. Dispersal via the MIOJet has been associated with the evolution of several taxa (Ragonieri et al. 2010; Muths et al. 2014) and our data and results fit this hypothesis well. Moreover, samples from the IAA, mostly pertaining to SUL and TAL, raise some interesting questions about *C. typus*' evolution, such as is there sympatry in Sarawak (ARC-TAL) or Aru (SUL-TAL), or is sympatry restricted to the Philippines? Do SUL and TAL exclude each other ecologically across the IAA? Is there philopatry in *C. typus* populations? To address these issues, further sampling is necessary, especially in the Philippines and the Pacific islands. Further marker sampling will also allow to test introgression between clades, which still cannot be ruled out.

Another important result of our study are questions raised about the monophyly and taxonomic status of *C. typus*. We conservatively identified three distinct clades (which could potentially be more since many locations were not sampled) and found evidence to treat each clade as a unique species. In addition, preliminary results indicate that these three clades may not be monophyletic, as TAL could be more closely related to *C. villadolidi* (though, as discussed, weakly supported; see Supplementary Fig. S2). Thus, *Caridina typus*' taxonomy is not clear, for instance, no type-locality was originally described, although it appears to be Mauritius (de Grave & Fransen 2011). This indicates that a full taxonomic revision is required.

Bibliography

- Adler, G. H. (1994). Avifaunal diversity and endemism on tropical Indian Ocean islands. *Journal of Biogeography*, 21(1), 85–95. <https://doi.org/10.2307/2845606>
- Agnarsson, I., Zambrana-Torrel, C. M., Flores-Saldana, N. P., & May-Collado, L. J. (2011). A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). *PLOS Currents Tree of Life*, 1. <https://doi.org/10.1371/currents.RRN1212>
- Ali, J. R., & Aitchison, J. C. (2008). Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). *Earth-Science Reviews*, 88(3–4), 145–166. <https://doi.org/10.1016/j.earscirev.2008.01.007>
- Allen, R., Najman, Y., Carter, A., Barfod, D., Bickle, M. J., Chapman, H. J., ... Parrish, R. R. (2008). Provenance of the Tertiary sedimentary rocks of the Indo-Burman Ranges, Burma (Myanmar): Burman arc or Himalayan-derived? *Journal of the Geological Society*, 165(6), 1045–1057. <https://doi.org/10.1144/0016-76492007-143>
- Almeida, F. C., Giannini, N. P., Simmons, N. B., & Helgen, K. M. (2014). Each flying fox on its own branch: A phylogenetic tree for *Pteropus* and related genera (Chiroptera: Pteropodidae). *Molecular Phylogenetics and Evolution*, 77, 83–95. <https://doi.org/10.1016/j.ympev.2014.03.009>
- Aneesh, S., & Sijikumar, S. (2018). Changes in the La Niña teleconnection to the Indian summer monsoon during recent period. *Journal of Atmospheric and Solar-Terrestrial Physics*, 167, 74–79. <https://doi.org/10.1016/j.jastp.2017.11.009>
- Ashalatha, B., Subrahmanyam, C., & Singh, R. N. (1991). Origin and compensation of Chagos-Laccadive ridge, Indian ocean, from admittance analysis of gravity and bathymetry data. *Earth and Planetary Science Letters*, 105, 47–54. [https://doi.org/10.1016/0012-821X\(91\)90119-3](https://doi.org/10.1016/0012-821X(91)90119-3)
- Ashwal, L. D., Wiedenbeck, M., & Torsvik, T. H. (2017). Archaean zircons in Miocene oceanic hotspot rocks establish ancient continental crust beneath Mauritius. *Nature Communications*, 8. <https://doi.org/10.1038/ncomms14086>
- Austin, J. J., Arnold, E. N., & Jones, C. G. (2004). Reconstructing an island radiation using ancient and recent DNA: the extinct and living day geckos (*Phelsuma*) of the Mascarene islands. *Molecular Phylogenetics and Evolution*, 31(1), 109–122. <https://doi.org/10.1016/j.ympev.2003.07.011>

- Avice, J. C. (2000). *Phylogeography: The History and Formation of Species*. Cambridge, MA: Harvard University Press.
- Avice, J. C. (2004). *Molecular Markers, Natural History, and Evolution* (2^a ed). Sunderland, MA: Sinauer Associates.
- Avice, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., ... Saunders, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, 18(1), 489–522. <https://doi.org/10.1146/annurev.es.18.110187.002421>
- Baele, G., Li, W. L. S., Drummond, A. J., Suchard, M. A., & Lemey, P. (2012). Accurate Model Selection of Relaxed Molecular Clocks in Bayesian Phylogenetics. *Molecular Biology and Evolution*, 30(2), 239–243. <https://doi.org/10.1093/molbev/mss243>
- Bănărescu, P. (1990). *Zoogeography of Freshwaters Vol. 1. General Distribution and Dispersal of Freshwater Animals*. Wiesbaden: Aula-Verlag.
- Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48.
- Barber-James, H. M. (2009). A preliminary phylogeny of Prosopistomatidae (Ephemeroptera) based on morphological characters of the larvae, and an assessment of their distribution. *Aquatic Insects*, 31(sup1), 149–166. <https://doi.org/10.1080/01650420903020502>
- Barron, E. J., & Peterson, W. H. (1991). The Cenozoic ocean circulation based on ocean General Circulation Model results. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 83(1), 1–28. [https://doi.org/10.1016/0031-0182\(91\)90073-Z](https://doi.org/10.1016/0031-0182(91)90073-Z)
- Beatty, J. (1986). The Synthesis and the Synthetic Theory. In W. Bechtel (Ed.), *Integrating Scientific Disciplines* (pp. 125–135). Springer Netherlands.
- Bernardes, S. C., Pepato, A. R., von Rintelen, T., von Rintelen, K., Page, T. J., Freitag, H., & de Bruyn, M. (2017). The complex evolutionary history and phylogeography of *Caridina typus* (Crustacea: Decapoda): long-distance dispersal and cryptic allopatric species. *Scientific Reports*, 7(1), 9044. <https://doi.org/10.1038/s41598-017-08494-w>
- Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., ... Das, I. (2006). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22(3), 148–155. <https://doi.org/10.1016/j.tree.2006.11.004>
- Biju, S. D., & Bossuyt, F. (2003) New frog family from India reveals an ancient biogeographical link with the Seychelles. *Nature*, 425, 711–714. <https://doi.org/10.1038/nature02019>

- Blaxter, M. (2003). Molecular systematics: Counting angels with DNA. *Nature*, *421*, 122–124. <https://doi.org/10.1038/421122a>
- Blaxter, M. L. (2004). The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 669–679. <https://doi.org/10.1098/rstb.2003.1447>
- Bohonak, A. J., & Jenkins, D. G. (2003). Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters*, *6*(8), 783–796. <https://doi.org/10.1046/j.1461-0248.2003.00486.x>
- Bossuyt, F., Meegaskumbura, M., Beenaerts, N., Gower, D. J., Pethiyagoda, R., Roelants, K., ... Milinkovitch, M. C. (2004). Local Endemism Within the Western Ghats-Sri Lanka Biodiversity Hotspot. *Science*, *306*(5695), 479–481. <https://doi.org/10.1126/science.1100167>
- Bouckaert, R., Alvarado-Mora, M. V., & Pinho, J. R. R. (2013). Evolutionary rates and HBV: issues of rate estimation with Bayesian molecular methods. *Antiviral Therapy*, *18*(3 Pt B), 497–503. <https://doi.org/10.3851/IMP2656>
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., ... Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology*, *10*(4), e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>
- Bouvier, E. L. (1925). Recherches Sur la morphologie, les variations la distribution géographique des crevettes de la famille des atyidae. *Encyclopedia Entomology (A)* *4*, 1–370.
- Braithwaite, C. J. R. (2016). The giant tortoise, Aldabrachelys, and its bearing on the biogeography and dispersal of terrestrial biota in the Western Indian Ocean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *461*, 449–459. <https://doi.org/10.1016/j.palaeo.2016.08.010>
- Briggs, J. C. (1989). The Historic Biogeography of India: Isolation or Contact? *Systematic Biology*, *38*(4), 322–332. <https://doi.org/10.2307/2992398>
- Brower, A. V. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences*, *91*(14), 6491–6495.
- Bruce, A. J. (1992). *Pycnisia raptor*, a new genus and species of predatory troglitic shrimp (Crustacea: Decapoda: Atyidae) from northern Australia. *Invertebrate Systematics*, *6*(3), 553–566. <https://doi.org/10.1071/IT9920553>

- Bybee, G. M., Ashwal, L. D., & Wilson, A. H. (2010). New evidence for a volcanic arc on the western margin of a rifting Rodinia from ultramafic intrusions in the Andriamena region, north-central Madagascar. *Earth and Planetary Science Letters*, 293(1), 42–53. <https://doi.org/10.1016/j.epsl.2010.02.017>
- Cai, Y., & Ng, P. K. L. (2001). The freshwater decapod crustaceans of halmahera, indonesia. *Journal of Crustacean Biology*, 21(3), 665–695. [https://doi.org/10.1651/0278-0372\(2001\)021\[0665:TFDCOH\]2.0.CO;2](https://doi.org/10.1651/0278-0372(2001)021[0665:TFDCOH]2.0.CO;2)
- Carranza, S., & Arnold, E. N. (2012). *A review of the geckos of the genus Hemidactylus (Squamata: Gekkonidae) from Oman based on morphology, mitochondrial and nuclear data, with descriptions of eight new species*. Auckland: Magnolia Press.
- Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, 22, 4369–4383. <https://doi.org/10.1111/mec.12413>
- Chace Jr., F. A. (1983). The *Atya*-like shrimps of the Indo-Pacific region (Decapoda, Atyidae). *Smithsonian Contribution to Zoology*, 384, 1–54.
- Chace Jr., F. A. (1997). The caridean shrimps (Crustacea: Decapoda) of the Albatross Philippine expedition 1907–1910. Part 7: Families Atyidae, Eugonatonotidae, Rhynchocinetidae, Bathypalaemonellidae, Processidae, and Hippolytidae. *Smithsonian Contributions to Zoology*, 587, 1–106.
- Chatterjee, S., & Scotese, C. R. (2010). The wandering Indian plate and its changing biogeography during the Late Cretaceous–Early Tertiary period. In: *New Aspects of Mesozoic Biodiversity* (Lecture Notes in Earth Sciences, Vol. 132). Ed. Bandyopadhyay, S., pp. 105–126. Berlin: Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-10311-7_7
- Chatterjee, S., Goswami, A., & Scotese, C. R. (2013). The longest voyage: Tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia. *Gondwana Research*, 23(1), 238–267. <https://doi.org/10.1016/j.gr.2012.07.001>
- Cheke, A. S., Pedrono, M., Bour, R., Anderson, A., Griffiths, C., Iverson, J. B., ... Walsh, M. (2017). Giant tortoises spread to western Indian Ocean islands by sea drift in pre-Holocene times, not by later human agency – response to Wilmé et al. (2016a). *Journal of Biogeography*, 44(6), 1426–1429. <https://doi.org/10.1111/jbi.12882>
- Chenoweth, S. F., & M. Hughes, J. (2003). Speciation and phylogeography in *Caridina indistincta*, a complex of freshwater shrimps from Australian heathland streams. *Marine and Freshwater Research*, 54(7), 807–812. <https://doi.org/10.1071/mf02168>

- Clemens, S., Prell, W., Murray, D., Shimmield, G., & Weedon, G. (1991). Forcing mechanisms of the Indian Ocean monsoon. *Nature*, 353, 720–725. <https://doi.org/10.1038/353720a0>
- Coffin, M. F., Pringle, M. S., Duncan, R. A., Gladchenko, T. P., Storey, M., Müller, R. D., & Gahagan, L. A. (2002). Kerguelen Hotspot magma output since 130 Ma. *Journal of Petrology*, 43(7), 1121–1137. <https://doi.org/10.1093/petrology/43.7.1121>
- Collier, J. S., Sansom, V., Ishizuka, O., Taylor, R. N., Minshull, T. A., & Whitmarsh, R. B. (2008). Age of Seychelles–India break-up. *Earth and Planetary Science Letters*, 272(1), 264–277. <https://doi.org/10.1016/j.epsl.2008.04.045>
- Collins, R. A., & Cruickshank, R. H. (2014). Known Knowns, Known Unknowns, Unknown Unknowns and Unknown Knowns in DNA Barcoding: A Comment on Downton et al. *Systematic Biology*, 0(0), 1–5. <https://doi.org/10.1093/sysbio/syu060>
- Conti, E., Eriksson, T., Schönenberger, J., Sytsma, K. J., Baum, D. A. (2002). Early Tertiary Out-of-India dispersal of Crypteroniaceae: evidence from phylogeny and molecular dating. *Evolution*, 56(10), 1931–1942. [https://doi.org/10.1554/0014-3820\(2002\)056\[1931:ETOOID\]2.0.CO;2](https://doi.org/10.1554/0014-3820(2002)056[1931:ETOOID]2.0.CO;2)
- Costello, M. J., Tsai, P., Wong, P. S., Cheung, A. K. L., Basher, Z., & Chaudhary, C. (2017). Marine biogeographic realms and species endemism. *Nature Communications*, 8, 1057. <https://doi.org/10.1038/s41467-017-01121-2>
- Covich, A. P., Palmer, M. A., & Crowl, T. A. (1999). The Role of Benthic Invertebrate Species in Freshwater Ecosystems: Zoobenthic species influence energy flows and nutrient cycling. *BioScience*, 49(2), 119–127. <https://doi.org/10.2307/1313537>
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8), 772–772. <https://doi.org/10.1038/nmeth.2109>
- Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle of life* (6th ed.). Reprint, Oxford: Oxford University Press.
- Datta-Roy, A., & Karanth, K. P. (2009). The Out-of-India hypothesis: What do molecules suggest? *Journal of Biosciences*, 34(5), 687–697. <https://doi.org/10.1007/s12038-009-0057-8>
- Daugherty, C. H., Cree, A., Hay, J. M., & Thompson, M. B. (1990). Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature*, 347, 177–179. <https://doi.org/10.1038/347177a0>

- Davie, P. J. F. (2002). Atyidae. In: *Crustacea: Malacostraca: Phyllocarida, Hoplocarida, Eucarida (Part 1)*. Eds. Wells, A., Houston, W. W. K., Davie, P. J. F., & Australian Biological Resources Study, pp 222–229. Melbourne: CSIRO Publishing.
- de Bruyn, M., Grail, W., & Carvalho, G. R. (2011). Anonymous nuclear markers for halfbeak fishes of the genus *Hemirhamphodon*. *Conservation Genetics Resources*, 3(1), 155–157. <https://doi.org/10.1007/s12686-010-9312-4>
- de Bruyn, M., Rüber, L., Nylinder, S., Stelbrink, B., Lovejoy, N. R., Lavoué, S., ... Carvalho, G. R. (2013). Paleo-Drainage Basin Connectivity Predicts Evolutionary Relationships across Three Southeast Asian Biodiversity Hotspots. *Systematic Biology*, 0(0), 1–13. <https://doi.org/10.1093/sysbio/syt007>
- de Grave, S. (2013). *Caridina typus*. The IUCN Red List of Threatened Species 2013: e.T198327A2520928. <https://dx.doi.org/10.2305/IUCN.UK.2013-1.RLTS.T198327A2520928.en>. Accessed at November 30th 2016.
- de Grave, S., & Fransen, C. H. J. M. (2011). Carideorum Catalogus: the recent species of the dendrobranchiate, stenopodidean, procarididean and caridean shrimps (Crustacea: Decapoda). *Zoologische Mededelingen*, 85, 195–589.
- de Queiroz, A. (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, 20(2), 68–73. <https://doi.org/10.1016/j.tree.2004.11.006>
- de Queiroz, K. (2007). Species Concepts and Species Delimitation. *Systematic Biology*, 56(6), 879–886. <https://doi.org/10.1080/10635150701701083>
- de Souza, M. L., & Moulton, T. P. (2005). The effects of shrimps on benthic material in a Brazilian island stream. *Freshwater Biology*, 50(4), 592–602. <https://doi:10.1111/j.1365-2427.2005.01348.x>
- de Wit, M. J. (2003). Madagascar: Heads it's a continent, tails it's an island. *Annual Review of Earth and Planetary Sciences*, 31, 213–248. <https://doi.org/10.1146/annurev.earth.31.100901.141337>
- Dobzhansky, T. (1937). *Genetics and the Origin of Species*. New York: Columbia University Press.
- 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. <https://doi.org/10.1093/molbev/mss075>
- Dunn, C. P. (2003). Keeping taxonomy based in morphology. *Trends in Ecology & Evolution*, 18(6), 270–271. [https://doi.org/10.1016/S0169-5347\(03\)00094-6](https://doi.org/10.1016/S0169-5347(03)00094-6)

- Emerson, B. C. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, *11*(6), 951–966. <https://doi.org/10.1046/j.1365-294X.2002.01507.x>
- Esson, J., Flower, M. F. J., Strong, D. F., & Upton, B. G. J. (1970). Geology of the Comores Archipelago, Western Indian Ocean. *Geological Magazine*, *107*(6), 549–557. <https://doi.org/10.1017/S0016756800058647>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491.
- 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. *Proceedings of the National Academy of Sciences*, *114*(29), E5864–E5870. <https://doi.org/10.1073/pnas.1704632114>
- Fenton, M. (2001). *Bats* (2nd ed.). Edition, New York, NY: Checkmark Books.
- Fièvet, E., & Eppe, R. (2002). Genetic differentiation among populations of the amphidromous shrimp *Atya innocous* (Herbst) and obstacles to their upstream migration. *Archiv für Hydrobiologie*, *153*, 287–300. <https://10.1127/archiv-hydrobiol/153/2002/287>
- Fluxus engineering. Available at: <https://www.fluxus-engineering.com/>. (accessed: 17.03.2016)
- Fransen, C. (2015). *Caridina* H. Milne Edwards, 1837 [in H. Milne Edwards, 1834-1840]. Accessed through: World Register of Marine Species at <https://www.marinespecies.org/aphia.php?p=taxdetails&id=240672> (accessed 02.10.2016).
- Frazão, A., Silva, H. R. da, & Russo, C. A. de M. (2015). The Gondwana Breakup and the History of the Atlantic and Indian Oceans Unveils Two New Clades for Early Neobatrachian Diversification. *PLOS ONE*, *10*(11), e0143926. <https://doi.org/10.1371/journal.pone.0143926>
- Fryer, G. (1977). Studies on the Functional Morphology and Ecology of the Atyid Prawns of Dominica. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *277*(952), 57–129. <https://10.1098/rstb.1977.0007>

- Fu, Y. X. (1997). Statistical Tests of Neutrality of Mutations against Population Growth, Hitchhiking and Background Selection. *Genetics*, 147(2), 915–925.
- Fujisawa, T., & Barraclough, T. G. (2013). Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. *Systematic Biology*, 62(5), 707–724. <https://doi.org/10.1093/sysbio/syt033>
- Fujita, J., Zenimoto, K., Iguchi, A., Kai, Y., Ueno, M., & Yamashita, Y. (2016). Comparative phylogeography to test for predictions of marine larval dispersal in three amphidromous shrimps. *Marine Ecology Progress Series*, 560, 105–120. <https://doi.org/doi:10.3354/meps11911>
- Gadgil, S., & Srinivasan, J. (2011). Seasonal prediction of the Indian monsoon. *Current Science*, 100(3), 343–353.
- Geller, J. B. (1999). Decline of a Native Mussel Masked by Sibling Species Invasion. *Conservation Biology*, 13(3), 661–664. <https://doi.org/10.1046/j.1523-1739.1999.97470.x>
- Gibb, G. C., & Penny, D. (2010). Two aspects along the continuum of pigeon evolution: A South-Pacific radiation and the relationship of pigeons within Neoaves. *Molecular Phylogenetics and Evolution*, 56(2), 698–706. <https://doi.org/10.1016/j.ympev.2010.04.016>
- Gourlan, A. T., Meynadier, L., & Allègre, C. J. (2008). Tectonically driven changes in the Indian Ocean circulation over the last 25 Ma: Neodymium isotope evidence. *Earth and Planetary Science Letters*, 267(1–2), 353–364. <https://doi.org/10.1016/j.epsl.2007.11.054>
- Gower, D. J., Kupfer, A., Oommen, O. J., Himstedt, W., Nussbaum, R. A., Loader, S. P., ... Wilkinson, M. (2002). A molecular phylogeny of ichthyophiid caecilians (Amphibia: Gymnophiona: Ichthyophiidae): out of India or out of South East Asia? *Proceedings of the Royal Society of London B: Biological Sciences*, 269, 1563–1569. <https://doi.org/10.1098/rspb.2002.2050>
- Grummer, J. A., Bryson, R. W., & Reeder, T. W. (2014). Species delimitation using Bayes Factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology*, 63(2), 119–133. <https://doi.org/10.1093/sysbio/syt069>

- Guillot, G., Mortier, F., & Estoup, A. (2005). Geneland: a computer package for landscape genetics. *Molecular Ecology Notes*, 5(3), 712–715. <https://doi.org/10.1111/j.1471-8286.2005.01031.x>
- Guillot, G., Renaud, S., Ledevin, R., Michaux, J., & Claude, J. (2012). A Unifying Model for the Analysis of Phenotypic, Genetic and Geographic Data. *Systematic Biology*, sys038. <https://doi.org/10.1093/sysbio/sys038>
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hall, M., & Keetley, J. (2009). *Otway Basin: stratigraphic and tectonic framework GeoScience Victoria 3D Victoria report 2*. Melbourne: Department of Primary Industries.
- Hall, R. (1996). Reconstructing Cenozoic SE Asia. *Geological Society, London, Special Publications*, 106(1), 153–184. <https://doi.org/10.1144/GSL.SP.1996.106.01.11>
- Hall, R. (2002). Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. *Journal of Asian Earth Sciences*, 20(4), 353–431. [https://doi.org/10.1016/S1367-9120\(01\)00069-4](https://doi.org/10.1016/S1367-9120(01)00069-4)
- Hall, R. (2011). Australia–SE Asia collision: plate tectonics and crustal flow. *Geological Society, London, Special Publications*, 355(1), 75–109. <https://doi.org/10.1144/SP355.5>
- Hall, R. (2012). Late Jurassic–Cenozoic reconstructions of the Indonesian region and the Indian Ocean. *Tectonophysics*, 570–571, 1–41. <https://doi.org/10.1016/j.tecto.2012.04.021>
- Hancock, M. A. (1998). The relationship between egg size and embryonic and larval development in the freshwater shrimp *Paratya australiensis* Kemp (Decapoda: Atyidae). *Freshwater Biology*, 39(4), 715–723. <https://10.1046/j.1365-2427.1998.00323.x>
- Hancock, M. A., Hughes, J. M., & Bunn, S. E. (1998). Influence of genetic and environmental factors on egg and clutch sizes among populations of *Paratya australiensis* Kemp (Decapoda: Atyidae) in upland rainforest streams, south-east Queensland. *Oecologia*, 115(4), 483–491. <https://10.1007/s004420050545>
- Hansen, D. M., Austin, J. J., Baxter, R. H., de Boer, E. J., Falcón, W., Norder, S. J., ... Warren, B. H. (2017). Origins of endemic island tortoises in the western Indian Ocean: a critique of the human-translocation hypothesis. *Journal of Biogeography*, 44(6), 1430–1435. <https://doi.org/10.1111/jbi.12893>

- Hawlitsek, O., Ramírez Garrido, S., & Glaw, F. (2017). How marine currents influenced the widespread natural overseas dispersal of reptiles in the Western Indian Ocean region. *Journal of Biogeography*, *44*(6), 1435–1440. <https://doi.org/10.1111/jbi.12940>
- Hawlitsek, O., Toussaint, E. F. A., Gehring, P.-S., Ratsavina, F. M., Cole, N., Crottini, A., ... Glaw, F. (2017). Gecko phylogeography in the Western Indian Ocean region: the oldest clade of *Ebenavia inunguis* lives on the youngest island. *Journal of Biogeography*, *44*(2), 409–420. <https://doi.org/10.1111/jbi.12912>
- Hay, J. M., Sarre, S. D., Lambert, D. M., Allendorf, F. W., & Daugherty, C. H. (2010). Genetic diversity and taxonomy: a reassessment of species designation in tuatara (*Sphenodon*). *Conservation Genetics*, *11*, 1063–1081. <https://doi.org/10.1007/s10592-009-9952-7>
- Hayashi, K. I., & Hamano, T. (1984). The complete larval development of *Caridina japonica* (Decapoda, Caridea, Atyidae) reared in the laboratory. *Zoological Science*, *1*, 571–589.
- Heads, M. (2005). Towards a panbiogeography of the seas. *Biological Journal of the Linnean Society*, *84*, 675–723. <https://doi.org/10.1111/j.1095-8312.2005.00466.x>
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*, 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Heinicke, M. P., Daza, J. D., Greenbaum, E., Jackman, T. R., & Bauer, A. M. (2014). Phylogeny, taxonomy and biogeography of a circum-Indian Ocean clade of leaf-toed geckos (Reptilia: Gekkota), with a description of two new genera. *Systematics and Biodiversity*, *12*(1), 23–42. <https://doi.org/10.1080/14772000.2013.877999>
- Heirtzler, J. R., Veevers, J. V., Bolli, H. M., Carter, A. N., Cook, P. J., Krasheninnikov, V. A., McKnight, B. K., ... Thayer, P. A. (1973). Age of the Floor of the Eastern Indian Ocean. *Science*, *180*(4089), 952–954. <https://doi.org/10.1126/science.180.4089.952>
- Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, *8*, 289. <https://doi.org/10.1186/1471-2148-8-289>
- Henderson, I. M. (1991). Biogeography without area? *Australian Systematic Botany*, *4*(1), 59–71. <https://doi.org/10.1071/SB9910059>
- Hillis, D. M., Moritz, C., & Mable, B. K. (1996). *Molecular systematics* (2nd ed.). Sunderland, MA: Sinauer Associates.
- Hocutt, C. H. (1987). Evolution of the Indian Ocean and the drift of India: A vicariant event. *Hydrobiologia*, *150*, 203–223. <https://doi.org/10.1007/BF00008704>

- Holder, M. T., Erdmann, M. V., Wilcox, T. P., Caldwell, R. L., & Hillis, D. M. (1999). Two living species of coelacanth? *Proceedings of the National Academy of Sciences*, 96(22), 12616–12620. <https://doi.org/10.1073/pnas.96.22.12616>
- Huelsenbeck, J. P., Larget, B., & Alfaro, M. E. (2004). Bayesian Phylogenetic Model Selection Using Reversible Jump Markov Chain Monte Carlo. *Molecular Biology and Evolution*, 21(6), 1123–1133. <https://doi.org/10.1093/molbev/msh123>
- Humphries, C.J., & Parenti, L.R. (2001). *Cladistic biogeography: interpreting patterns of plant and animal distributions* (2nd ed.). Oxford: Oxford University Press.
- Hurwood, D. A., & Hughes, J. M. (2001). Nested clade analysis of the freshwater shrimp, *Caridina zebra* (Decapoda: Atyidae), from north-eastern Australia. *Molecular Ecology*, 10(1), 113–125. <https://doi.org/10.1046/j.1365-294X.2001.01175.x>
- Huxley, T. H. (1880). *The Crayfish – an introduction to the study of biology*. Reprint, Cambridge, MA: MIT Press.
- Jalihai, D. R., Almelkar, G. B., & Sankolli, K. N. (1994). Atyid Shrimps of the Genus *Caridina* H. Milne Edwards, 1837. Potential Crustacean Material for Experimental Biology. *Crustaceana*, 66(2), 178–183.
- Johnson, D. S. (1960). Some aspects of the distribution of freshwater organisms in Indo-Pacific area, and their relevance to the validity of the concepts of an Oriental region in zoogeography. *Proceedings of the Centennial and Biocentennial Congress of Biology*, Singapore, pp. 170–181.
- Johnson, D. S. (1961). Notes on the freshwater crustaceans of Malaya. 1. The Atyidae. *Bulletin Raffles Museum Singapore* 26, 120–125.
- Johnson, D. S. (1963). Distributional and other notes on some fresh-water prawns (Atyidae and Palaemonidae) mainly from the Indo-West Pacific region. *Bulletin of the National Museum*, 32, 5–30.
- Jones, G. (2016). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, 1–21. <https://doi.org/10.1007/s00285-016-1034-0>
- Jones, G., Aydin, Z., & Oxelman, B. (2014). DISSECT: an assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, btu770. <https://doi.org/10.1093/bioinformatics/btu770>
- Jurado-Rivera, J. A., Jaume, D., Juan, C., & Pons, J. (2016). The complete mitochondrial genome of the cave shrimp *Typhlatya miravetensis* (Decapoda, Atyidae) and its

- systematic position. *Mitochondrial DNA Part B*, 1(1), 847–848. <https://doi.org/10.1080/23802359.2016.1238756>
- Jurado-Rivera, J. A., Pons, J., Alvarez, F., Botello, A., Humphreys, W. F., Page, T. J., ... Jaume, D. (2017). Phylogenetic evidence that both ancient vicariance and dispersal have contributed to the biogeographic patterns of anchialine cave shrimps. *Scientific Reports*, 7(1), 2852. <https://doi.org/10.1038/s41598-017-03107-y>
- 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. <https://doi.org/10.1093/molbev/mst010>
- Keller, G. (2014). Deccan volcanism, the Chicxulub impact, and the end-Cretaceous mass extinction: Coincidence? Cause and effect? In: *Volcanism, Impacts, and Mass Extinctions: Causes and Effects*. Eds. Keller, G., & Kerr, A. C., pp. 57–74. McLean, VA: Geological Society of America Special Paper.
- Keller, G., Bhowmick, P. K., Upadhyay, H., Dave, A., Reddy, A. N., Jaiprakash, B. C., & Adatte, T. (2011). Deccan volcanism linked to the Cretaceous-Tertiary boundary mass extinction: New evidence from ONGC wells in the Krishna-Godavari basin. *Journal Geological Society Of India*, 78 (5), 399–428. <https://doi.org/10.1007/s12594-011-0107-3>
- Khosla, A. (2014). Upper Cretaceous (Maastrichtian) charophyte gyrogonites from the Lameta Formation of Jabalpur, Central India: Palaeobiogeographic and palaeoecological implications. *Acta Geologica Polonica*, 64, 311–323. <https://doi.org/10.2478/agp-2014-0016>
- Klaus, S., Schubart, C. D., & Brandis, D. (2006). Phylogeny, biogeography and a new taxonomy for the Gecarcinucoidea Rathbun, 1904 (Decapoda: Brachyura). *Organisms Diversity & Evolution*, 6(3), 199–217. <https://doi.org/10.1016/j.ode.2005.09.006>
- Knapp, S. (2005). Biogeography – space, form and time. *Journal of Biogeography*, 32(1), 3–4. <https://doi.org/10.1111/j.1365-2699.2004.01229.x>
- Knowlton, N., Weigt, L. A., Solorzano, L. A., Mills, D. K., & Bermingham, E. (1993). Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the isthmus of Panama. *Science*, 260(5114), 1629–1632. <https://doi.org/10.1126/science.8503007>
- Krause, D.W. (2003). Late Cretaceous vertebrates of Madagascar: a window into Gondwanan biogeography at the end of the age of dinosaurs. In: *The Natural History of Madagascar*.

- Eds. Goodman, S. M. and Benstead, J. P., pp. 40–47. Chicago: University of Chicago Press.
- Krause, D. W. (2010). Washed up in Madagascar. *Nature*, 463, 613–614. <https://doi.org/10.1038/463613a>
- Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian Analysis of Biogeography when the Number of Areas is Large. *Systematic Biology*, 1–16. <https://doi.org/10.1093/sysbio/syt040>
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., ... Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>
- Latif, M., Hannachi, A., & Syed, F. S. (2017). Analysis of rainfall trends over Indo-Pakistan summer monsoon and related dynamics based on CMIP5 climate model simulations. *International Journal of Climatology*, 38(1). <https://doi.org/10.1002/joc.5391>
- Lemey, P., Rambaut, A., Drummond, A. J., & Suchard, M. A. (2009). Bayesian Phylogeography Finds Its Roots. *PLOS Computational Biology*, 5(9), e1000520. <https://doi.org/10.1371/journal.pcbi.1000520>
- Linder, H. P. (2001). On areas of endemism, with an example from the African Restionaceae. *Systematic Biology*, 50(6), 892–912. <https://doi.org/10.1080/106351501753462867>
- Lipscomb, D., Platnick, N., & Wheeler, Q. (2003). The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology & Evolution*, 18(2), 65–66. [https://doi.org/10.1016/S0169-5347\(02\)00060-5](https://doi.org/10.1016/S0169-5347(02)00060-5)
- Lohman, D. J., de Bruyn, M., Page, T., Rintelen, K. von, Hall, R., Ng, P. K. L., ... Rintelen, T. von. (2011). Biogeography of the Indo-Australian Archipelago. *Annual Review of Ecology, Evolution, and Systematics*, 42(1), 205–226. <https://doi.org/10.1146/annurev-ecolsys-102710-145001>
- Lopez-Vaamonde, C., Godfray, H. C. J., West, S. A., Hansson, C., & Cook, J. M. (2005). The evolution of host use and unusual reproductive strategies in *Achrysocharoides* parasitoid wasps. *Journal of Evolutionary Biology*, 18(4), 1029–1041. <https://doi.org/10.1111/j.1420-9101.2005.00900.x>
- March, J. G., & Pringle, C. M. (2003). Food Web Structure and Basal Resource Utilization along a Tropical Island Stream Continuum, Puerto Rico. *Biotropica*, 35(1), 84–93. <https://doi.org/10.1111/j.1744-7429.2003.tb00265.x>
- Marshall, D. C., Hill, K. B. R., Moulds, M., Vanderpool, D., Cooley, J. R., Mohagan, A. B., & Simon, C. (2016). Inflation of Molecular Clock Rates and Dates: Molecular

- Phylogenetics, Biogeography, and Diversification of a Global Cicada Radiation from Australasia (Hemiptera: Cicadidae: Cicadettini). *Systematic Biology*, 65(1), 16–34. <https://doi.org/10.1093/sysbio/syv069>
- Matzke, N. J. (2013). BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis. Accessed through: R scripts, CRAN: The Comprehensive R Archive Network <http://CRAN.R-project.org/package=BioGeoBEARS> (accessed 02.10.2017).
- McDougall, I., & Chamalaun, F. H. (1969). Isotopic Dating and Geomagnetic Polarity Studies on Volcanic Rocks from Mauritius, Indian Ocean. *Geological Society of America Bulletin*, 80(8), 1419–1442. [https://doi.org/10.1130/0016-7606\(1969\)80\[1419:IDAGPS\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1969)80[1419:IDAGPS]2.0.CO;2)
- McLean, D. W. (1978). A Terminal Mesozoic "Greenhouse": Lessons from the Past. *Science*, 4(4354), 401–406. <https://10.1126/science.201.4354.401>
- Metcalf, I. (2013). Gondwana dispersion and Asian accretion: Tectonic and palaeogeographic evolution of eastern Tethys. *Journal of Asian Earth Sciences*, 66, 1–33. <https://doi.org/10.1016/j.jseaes.2012.12.020>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Gateway Computing Environments Workshop (GCE), 2010* (pp. 1–8). <https://doi.org/10.1109/GCE.2010.5676129>
- Mitchell, K. J., Wood, J. R., Scofield, R. P., Llamas, B., & Cooper, A. (2014). Ancient mitochondrial genome reveals unsuspected taxonomic affinity of the extinct Chatham duck (*Pachyanas chathamica*) and resolves divergence times for New Zealand and sub-Antarctic brown teals. *Molecular Phylogenetics and Evolution*, 70, 420–428. <https://doi.org/10.1016/j.ympev.2013.08.017>
- Morley, R. J. (2000). *Origin and Evolution of Tropical Rain Forests*. Chichester, UK: JohnWiley & Sons.
- Moyle, R. G., Andersen, M. J., Oliveros, C. H., Steinheimer, F. D., & Reddy, S. (2012). Phylogeny and Biogeography of the Core Babblers (Aves: Timaliidae). *Systematic Biology*, 61(4), 631–651. <https://doi.org/10.1093/sysbio/sys027>
- Muths, D., Tessier, E., & Bourjea, J. (2014). Genetic structure of the reef grouper *Epinephelus merra* in the West Indian Ocean appears congruent with biogeographic and oceanographic boundaries. *Marine Ecology*, 36(3), 447–461. <https://doi.org/10.1111/maec.12153>

- Nelson, G., & Platnick, N. I. (1981). *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia University Press.
- Noguera-urbano, E. A. (2016) Areas of endemism: travelling through space and the unexplored dimension. *Systematics and Biodiversity*, 14(2), 131–139. <https://doi.org/10.1080/14772000.2015.1135196>
- Noonan, B. P., & Chippindale, P. T. (2006). Dispersal and vicariance: The complex evolutionary history of boid snakes. *Molecular Phylogenetics and Evolution*, 40(2), 347–358. <https://doi.org/10.1016/j.ympev.2006.03.010>
- Ortmann, A. E. (1894). A Study of the Systematic and Geographical Distribution of the Decapod Family Atyidae Kingsley. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 46, 397–416.
- Page, T. J., & Hughes, J. M. (2007). Radically different scales of phylogeographic structuring within cryptic species of freshwater shrimp (Atyidae: Caridina). *Limnology and Oceanography*, 52(3), 1055–1066. <https://doi.org/10.4319/lo.2007.52.3.1055>
- Page, T. J., Choy, S. C., & Hughes, J. M. (2005). The taxonomic feedback loop: symbiosis of morphology and molecules. *Biology Letters*, 1, 139–142. <https://doi.org/10.1098/rsbl.2005.0298>
- Page, T. J., Torati, L. S., Cook, B. D., Binderup, A., Pringle, C. M., Reuschel, S., ... Hughes, J. M. (2013). Invertébrés Sans Frontières: Large Scales of Connectivity of Selected Freshwater Species among Caribbean Islands. *Biotropica*, 45(2), 236–244. <https://doi.org/10.1111/j.1744-7429.2012.00900.x>
- Page, T. J., von Rintelen, K., & Hughes, J. M. (2007). An island in the stream: Australia's place in the cosmopolitan world of Indo-West Pacific freshwater shrimp (Decapoda: Atyidae: Caridina). *Molecular Phylogenetics and Evolution*, 43, 645–659. <https://doi.org/10.1016/j.ympev.2006.08.007>
- Patterson, C., & Owen, H. G. (1991). Indian Isolation or Contact? A Response to Briggs. *Systematic Biology*, 40(1), 96–100. <https://doi.org/10.2307/2992225>
- Pereira, S. L., Johnson, K. P., Clayton, D. H., Baker, A. J., & Paterson, A. (2007). Mitochondrial and Nuclear DNA Sequences Support a Cretaceous Origin of Columbiformes and a Dispersal-Driven Radiation in the Paleogene. *Systematic Biology*, 56(4), 656–672. <https://doi.org/10.1080/10635150701549672>
- Pillon, Y., & Buerki, S. (2017). How old are island endemics? *Biological Journal of the Linnean Society*, 121(2), 469–474. <https://doi.org/10.1093/biolinnean/blx005>

- Plummer, Ph. S., & Belle, E. R. (1995). Mesozoic tectono-stratigraphic evolution of the Seychelles microcontinent. *Sedimentary Geology*, 96(1–2), 73–91. [https://doi.org/10.1016/0037-0738\(94\)00127-G](https://doi.org/10.1016/0037-0738(94)00127-G)
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Systematic Biology*, 55(4), 595–609. <https://doi.org/10.1080/10635150600852011>
- Porter, M. L., Pérez-Losada, M., & Crandall, K. A. (2005). Model-based multi-locus estimation of decapod phylogeny and divergence times. *Molecular Phylogenetics and Evolution*, 37, 355–369. <https://doi.org/10.1016/j.ympev.2005.06.021>
- Prasad, G. V. R. (2012) Vertebrate biodiversity of the Deccan volcanic province of India: A review. *Bulletin de la Société Géologique de France*, 183(6), 597–610. <https://doi.org/10.2113/gssgfbull.183.6.597>
- Proudlove, G., & Wood, P. J. (2003). The blind leading the blind: cryptic subterranean species and DNA taxonomy. *Trends in Ecology & Evolution*, 18(6), 272–273. [https://doi.org/10.1016/S0169-5347\(03\)00095-8](https://doi.org/10.1016/S0169-5347(03)00095-8)
- Provine, W. (1988). Progress in evolution and meaning in life. In: *Evolutionary progress*. Ed. Nitecki, M. H, pp. 49–74. Chicago: University of Chicago Press.
- Pyron, R. A. (2014). Biogeographic Analysis Reveals Ancient Continental Vicariance and Recent Oceanic Dispersal in Amphibians. *Systematic Biology*, 63(5), 779–797. <https://doi.org/10.1093/sysbio/syu042>
- Raftery, A., Newton, M., Satagopan, J., Krivitsky, P., Bernardo, J. M., Bayarri, M. J., & Berger, J. O. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In: *Bayesian Statistics 8*. Eds. Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A. F. M., & West, M., pp. 1–45. Oxford: Oxford University Press.
- Rage, J.-C. (1996). Le peuplement animal de Madagascar: une composante venue de Laurasia est-elle envisageable? In: *Biogéographie de Madagascar*. Ed. Lourenço, W. R., pp. 27–35. Paris: ORSTOM.
- Rage, J.-C. (2003). Relationships of the Malagasy fauna during the Late Cretaceous: Northern or southern routes? *Acta Palaeontologica Polonica*, 48(4), 661–662.
- Ragionieri, L., Cannicci, S., Schubart, C. D., & Fratini, S. (2010). Gene flow and demographic history of the mangrove crab *Neosarmatium meinerti*: A case study from the western

- Indian Ocean. *Estuarine, Coastal and Shelf Science*, 86(2), 179–188. <https://doi.org/10.1016/j.ecss.2009.11.002>
- Raval, U., & Veeraswamy, K. (2003). India-Madagascar Separation: Breakup Along a Pre-existing Mobile Belt and Chipping of the Craton. *Gondwana Research*, 6(3), 467–485. [https://doi.org/10.1016/S1342-937X\(05\)70999-0](https://doi.org/10.1016/S1342-937X(05)70999-0)
- Raxworthy, C. J., Forstner, M. R. J., & Nussbaum, R. A. (2002). Chameleon radiation by oceanic dispersal. *Nature*, 415(6873), 784–787. <https://doi.org/10.1038/415784a>
- 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. <https://doi.org/10.1080/10635150701883881>
- Reeves, D., & de Wit, M. (2000). Making ends meet in Gondwana: Retracing the transforms of the Indian Ocean and reconnecting continental shear zones. *Terra Nova*, 12(6), 272–280. <https://doi.org/10.1046/j.1365-3121.2000.00309.x>
- 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, 196. <https://doi.org/10.1186/1471-2148-12-196>
- Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., ... Pandolfi, J. M. (2008). Hopping Hotspots: Global Shifts in Marine Biodiversity. *Science*, 321(5889), 654–657. <https://doi.org/10.1126/science.1155674>
- Renne, P. R., Sprain, C. J., Richards, M. A., Self, S., Vanderkluisen, L., & Pande, K. (2015). State shift in Deccan volcanism at the Cretaceous-Paleogene boundary, possibly induced by impact. *Science*, 350(6256), 76–78. <https://doi.org/10.1126/science.aac7549>
- Reyda, F. B., & Olson, P. D. (2003). Cestodes of cestodes of peruvian freshwater stingrays. *Journal of Parasitology*, 89(5), 1018–1024. <https://doi.org/10.1645/GE-3143>
- Riek, E. F. (1953). The Australian freshwater prawns of the family Atyidae. *Records of the Australian Museum*, 23(3), 101–121. <https://doi.org/doi:10.3853/j.0067-1975.23.1953.625>
- Rocha, S., Carretero, M. A., Vences, M., Glaw, F., & James Harris, D. (2006). Deciphering patterns of transoceanic dispersal: the evolutionary origin and biogeography of coastal lizards (*Cryptoblepharus*) in the Western Indian Ocean region. *Journal of Biogeography*, 33(1), 13–22. <https://doi.org/10.1111/j.1365-2699.2005.01375.x>
- Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>

- Roxy, M. K., Ritika, K., Terray, P., & Masson, S. (2014). The Curious Case of Indian Ocean Warming. *Journal of Climate*, 27(22), 8501–8509. <https://doi.org/10.1175/JCLI-D-14-00471.1>
- Roxy, M. K., Ritika, K., Terray, P., Murtugudde, R., Ashok, K., & Goswami, B. N. (2015). Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient. *Nature Communications*, 6. <https://doi.org/10.1038/ncomms8423>
- Sahni, A. (1984). Cretaceous-Paleocene terrestrial faunas of India: Lack of endemism during drifting of the Indian Plate. *Science*, 226(4673), 441–443. <https://doi.org/10.1126/science.226.4673.441>
- Sampson, S. D., Witmer, L. M., Forster, C. A., Krause, D. W., O'Connor, P. M., Dodson, P., & Ravoavy, F. (1998). Predatory dinosaur remains from Madagascar: Implications for the Cretaceous biogeography of Gondwana. *Science*, 280(5366), 1048–1051. <https://doi.org/10.1126/science.280.5366.1048>
- Satler, J. D., Carstens, B. C., & Hedin, M. (2013). Multilocus Species Delimitation in a Complex of Morphologically Conserved Trapdoor Spiders (Mygalomorphae, Antrodiaetidae, Aliatypus). *Systematic Biology*, 62(6), 805–823. <https://doi.org/10.1093/sysbio/syt041>
- Schatz, G. E. (1996). Malagasy/Indo-Australo-Malesian phytogeographic connections. In: *Biogéographie de Madagascar*. Ed. Lourenço, W. R., pp. 73–83. Paris: ORSTOM.
- Schoene, B., Samperton, K. M., Eddy, M. P., Keller, G., Adatte, T., Bowring, S. A., ... Gertsch, B. (2015). U-Pb geochronology of the Deccan Traps and relation to the end-Cretaceous mass extinction. *Science*, 347(6218), 182–184. <https://doi.org/10.1126/science.aaa0118>
- Schott, F. A., Dengler, M., & Schoenefeldt, R. (2002). The shallow overturning circulation of the Indian Ocean. *Progress in Oceanography*, 53(1), 57–103. [https://doi.org/10.1016/S0079-6611\(02\)00039-3](https://doi.org/10.1016/S0079-6611(02)00039-3)
- Schulte, P., Alegret, L., Arenillas, I., Arz, J. A., Barton, P. J., Brown, P. R., ... Willumsen, P. S. (2010). The Chicxulub Asteroid Impact and Mass Extinction at the Cretaceous-Paleogene Boundary. *Science*, 327(5970), 1214–1218. <https://doi.org/10.1126/science.1177265>
- Scotland, R., Hughes, C., Bailey, D., & Wortley, A. (2003). The Big Machine and the much-maligned taxonomist. *Systematics and Biodiversity*, 1(2), 139–143. <https://doi.org/10.1017/S1477200003001178>

- Shapiro, B., Sibthorpe, D., Rambaut, A., Austin, J., Wragg, G. M., Bininda-Emonds, O. R. P., ... Cooper, A. (2002). Flight of the Dodo. *Science*, 295(5560), 1683–1683. <https://doi.org/10.1126/science.295.5560.1683>
- Short, J. W., & Doumenq, E. (2003). Atyidae and Palaemonidae, Freshwater Shrimps. In: *The Natural History of Madagascar*. Eds. Goodman, S. M., & Benstead, J. P., pp. 603–608. Chicago: University of Chicago Press.
- Shy, J.-Y., Lai, H.-T., & Yu, H.-P. (2001). On the Larval Development of *Caridina formosae* (Decapoda, Atyidae) reared in the Laboratory. *Crustaceana*, 74(10), 1159–1168.
- Simpson, G. G. (1952). Probabilities of dispersal in geologic time. *Bulletin of the American Museum of Natural History*, 99, 163–176.
- Springer, V. G. (1988). The Indo-Pacific blennioid fish genus *Ecsenius*. *Smithsonian Contributions to Zoology*, 465, 1–160.
- Springer, V. G. (1999). Are the Indonesian and western Indian Ocean coelacanths conspecific: a prediction. *Environmental Biology of Fishes*, 54, 453–456. <https://doi.org/10.1023/A:1007524929593>
- Streelman, J. T., & Danley, P. D. (2003). The stages of vertebrate evolutionary radiation. *Trends in Ecology & Evolution*, 18(3), 126–131. [https://doi.org/10.1016/S0169-5347\(02\)00036-8](https://doi.org/10.1016/S0169-5347(02)00036-8)
- Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, 114(7), 1607–1612. <https://doi.org/10.1073/pnas.1607921114>
- Suzuki, H., Tanigawa, N., Nagatomo, T., & Tsuda, E. (1993). Distribution of freshwater caridean shrimps and prawns (Atyidae and Palaemonidae) from Southern Kyushu and adjacent islands, Kagoshima Prefecture, Japan. *Crustacean Research*, 22, 55–64. https://doi.org/10.18353/crustacea.22.0_55
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30(12), 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., & Vogler, A. P. (2003). A plea for DNA taxonomy. *Trends in Ecology & Evolution*, 18(2), 70–74. [https://doi.org/10.1016/S0169-5347\(02\)00041-1](https://doi.org/10.1016/S0169-5347(02)00041-1)
- Thomas, D. J., Bralower, T. J., & Jones, C. E. (2003). Neodymium isotopic reconstruction of late Paleocene–early Eocene thermohaline circulation. *Earth and Planetary Science Letters*, 209(3–4), 309–322. [https://doi.org/10.1016/S0012-821X\(03\)00096-7](https://doi.org/10.1016/S0012-821X(03)00096-7)

- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673–4680. <https://doi.org/10.1093/nar/22.22.4673>
- Thornton, I. W. B., Runciman, D., Cook, S., Lumsden, L. F., Partomihardjo, T., Schedvid, N. K., ... Ward, S. A. (2002). How important were stepping stones in the colonization of Krakatau? *Biological Journal of the Linnean Society*, 77(3), 275–317. <https://doi.org/10.1046/j.1095-8312.2002.00113.x>
- Tolley, K. A., Townsend, T. M., & Vences, M. (2013). Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1759), 20130184. <https://doi.org/10.1098/rspb.2013.0184>
- Tomczak, M., & Godfrey, J. S. (1994). *Regional oceanography: an introduction*. Tarrytown, New York: Permagon.
- Torsvik, T. H., Amundsen, H., Hartz, E. H., Corfu, F., Kuszniir, N., Gaina, C., ... Jamtveit, B. (2013). A Precambrian microcontinent in the Indian Ocean. *Nature Geoscience*, 6(3), 223–227. <https://doi.org/10.1038/ngeo1736>
- Torsvik, T. H., Ashwal, L. D., Tucker, R. D., & Eide, E. A. (2001). Neoproterozoic geochronology and palaeogeography of the Seychelles microcontinent: the India link. *Precambrian Research*, 110(1), 47–59. [https://doi.org/10.1016/S0301-9268\(01\)00180-2](https://doi.org/10.1016/S0301-9268(01)00180-2)
- Trénel, P., Gustafsson, M. H. G., Baker, W. J., Asmussen-Lange, C. B., Dransfield, J., & Borchsenius, F. (2007). Mid-Tertiary dispersal, not Gondwanan vicariance explains distribution patterns in the wax palm subfamily (Ceroxyloideae: Areaceae). *Molecular Phylogenetics and Evolution*, 45(1), 272–288. <https://doi.org/10.1016/j.ympev.2007.03.018>
- van Oosterzee, P. (1997). *Where Worlds Collide: The Wallace Line*. Ithaca, NY: Cornell University Press.
- van Steenis, C. G. G. J. (1962). The land-bridge theory in botany, with particular reference to tropical plants. *Blumea*, 11(2), 235–542.
- Vences, M., Wollenberg, K. C., Vieites, D. R., & Lees, D. C. Madagascar as a model region of species diversification. *Trends in Ecology & Evolution*, 24(8), 456–465. <https://dx.doi.org/10.1016/j.tree.2009.03.011>

- Venkatesan, T. R., Pande, K., & Gopalan K. (1993). Did Deccan volcanism pre-date the Cretaceous/Tertiary transition? *Earth and Planetary Science Letters*, *119*(1–2), 181–189. [https://doi.org/10.1016/0012-821X\(93\)90015-2](https://doi.org/10.1016/0012-821X(93)90015-2)
- Verma, O., Khosla, A., Goin, F. J., & Kaur, J. (2016). Historical biogeography of the Late Cretaceous vertebrates of India: Comparison of geophysical and paleontological data. *New Mexico Museum of Natural History and Science Bulletin*, *71*, 317–330
- Verzhbitsky, E. V. (2003). Geothermal regime and genesis of the Ninety-East and Chagos-Laccadive ridges. *Journal of Geodynamics*, *35*(3), 289–302. [https://doi.org/10.1016/S0264-3707\(02\)00068-6](https://doi.org/10.1016/S0264-3707(02)00068-6)
- von Rintelen, K., Glaubrecht, M., Schubart, C. D., Wessel, A., & Von Rintelen, T. (2010). Adaptive Radiation and Ecological Diversification of Sulawesi's Ancient Lake Shrimps. *Evolution*, *64*(11), 3287–3299. <https://doi.org/10.1111/j.1558-5646.2010.01043.x>
- von Rintelen, K., Page, T. J., Cai, Y., Roe, K., Stelbrink, B., Kuhajda, B. R., Iliffe, T. M., Hughes, J., & von Rintelen, T. (2012). Drawn to the dark side: A molecular phylogeny of freshwater shrimps (Crustacea: Decapoda: Caridea: Atyidae) reveals frequent cave invasions and challenges current taxonomic hypotheses. *Molecular Phylogenetics and Evolution*, *63*(1), 82–96. <https://doi.org/10.1016/j.ympev.2011.12.015>
- von Rintelen, K., von Rintelen, T., & Glaubrecht, M. (2007). Molecular phylogeny and diversification of freshwater shrimps (Decapoda, Atyidae, Caridina) from ancient Lake Poso (Sulawesi, Indonesia)—The importance of being colourful. *Molecular Phylogenetics and Evolution*, *45*(3), 1033–1041. <https://doi.org/10.1016/j.ympev.2007.07.002>
- Wallace, A. R. (1881). *Island Life*. New York: Harper & Bros.
- Warren, B. H., Strasberg, D., Bruggemann, J. H., Prys-Jones, R. P., & Thébaud, C. (2010). Why does the biota of the Madagascar region have such a strong Asiatic flavour? *Cladistics*, *26*(5), 526–538. <https://doi.org/10.1111/j.1096-0031.2009.00300.x>
- Whatley, R. C. (2012). The 'Out of India' hypothesis: Further supporting evidence from the extensive endemism of Maastrichtian non-marine ostracoda from the Deccan volcanic region of peninsular India. *Revue de Paléobiologie*, *11*, 229–248.
- Wikström, N., Avino, M., Razafimandimbison, S. G., & Bremer, B. (2010). Historical biogeography of the coffee family (Rubiaceae, Gentianales) in Madagascar: case studies from the tribes Knoxieae, Naucleae, Paederieae and Vanguerieae. *Journal of Biogeography*, *37*(6), 1094–1113. <https://doi.org/10.1111/j.1365-2699.2009.02258.x>

- Will, K. W., & Rubinoff, D. (2004). Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics*, 20(1), 47–55. <https://doi.org/10.1111/j.1096-0031.2003.00008.x>
- Williams, A. D. (1980). *Australian freshwater life: the invertebrates of Australian inland water* (2^a ed). South Yarra: Macmillan.
- Woolschot, L., Hughes, J. M., & Bunn, S. E. (1999). Dispersal among populations of *Caridina* sp. (Decapoda: Atyidae) in coastal lowland streams, south-eastern Queensland, Australia. *Marine and Freshwater Research*, 50(7), 681–688. <https://doi.org/10.1071/MF99011>
- Yam, R. S. W., & Dudgeon, D. (2005). Stable isotope investigation of food use by *Caridina* spp. (Decapoda: Atyidae) in Hong Kong streams. *Journal of the North American Benthological Society*, 24(1), 68–81. [https://doi.org/10.1899/0887-3593\(2005\)024<0068:SIIOFU>2.0.CO;2](https://doi.org/10.1899/0887-3593(2005)024<0068:SIIOFU>2.0.CO;2)
- Yam, R. S. W., & Dudgeon, D. (2006). Production dynamics and growth of atyid shrimps (Decapoda: *Caridina* spp.) in 4 Hong Kong streams: the effects of site, season, and species. *Journal of the North American Benthological Society*, 25(2), 406–416. [https://doi.org/10.1899/0887-3593\(2006\)25\[406:PDAGOA\]2.0.CO;2](https://doi.org/10.1899/0887-3593(2006)25[406:PDAGOA]2.0.CO;2)
- Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, 61(5), 854–865. <https://doi.org/10.1093/czoolo/61.5.854>
- Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences*, 107(20), 9264–9269. <https://doi.org/10.1073/pnas.0913022107>
- Yoder, A. D., & Nowak, M. D. (2006). Has Vicariance or Dispersal Been the Predominant Biogeographic Force in Madagascar? Only Time Will Tell. *Annual Review of Ecology, Evolution, and Systematics*, 37(1), 405–431. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110239>
- Yu, Y., Harris, A. J., & He, X. (2010). S-DIVA (Statistical Dispersal-Vicariance Analysis): A tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, 56, 848–850. <https://doi.org/10.1016/j.ympev.2010.04.011>
- 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. <https://doi.org/10.1016/j.ympev.2015.03.008>

- Zink, R. M., Blackwell-Rago, R. C., & Ronquist, F. (2000). The shifting roles of dispersal and vicariance in biogeography. *Proceedings of the Royal Society of London B: Biological Sciences*, 267, 497–503. <https://doi.org/10.1098/rspb.2000.1028>
- Zuckerkandl, E., & Pauling, L. (1965). Molecules as Documents of Evolutionary History. *Journal of Theoretical Biology*, 8, 357–366. [https://doi.org/10.1016/0022-5193\(65\)90083-4](https://doi.org/10.1016/0022-5193(65)90083-4)
- Zwickl, D. J. (2006). *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion* (PhD). The University of Texas, Austin, TX. Retrieved from <https://search.proquest.com/docview/304984400>

Appendix

Table S1: Samples used in this study

Taxon	Location	Sample ID	Coordinates	GenBank accession numbers						
				16S	COI	Ct33	Ct51	28S		
<i>Caridina typus</i>	Indonesia	Aru								
		CT_Ar01	06° 10' 42.53" S	134° 18' 04.30" E	KY069330	KY069449	KY069546	KY069642	KY069643	KY069719
		CT_Ar02	06° 10' 42.53" S	134° 18' 04.30" E	KY069331	KY069450	KY069547	KY069644	KY069645	KY069720
		CT_Ar03	06° 10' 42.53" S	134° 18' 04.30" E	KY069332	KY069451	KY069548	KY069646	KY069647	KY069721
		CT_Ar04	06° 10' 42.53" S	134° 18' 04.30" E	KY069333	KY069452	KY069549	KY069648	KY069649	KY069722
		CT_Ar05	06° 10' 42.53" S	134° 18' 04.30" E	KY069334	KY069453	KY069550	KY069650	KY069651	KY069723
		CT_Ar06	06° 10' 42.53" S	134° 18' 04.30" E	KY069335	KY069454	KY069551	KY069652	KY069653	KY069724
		CT_Ar07	06° 10' 42.53" S	134° 18' 04.30" E	KY069336	KY069455	KY069552	KY069654	KY069655	KY069725
		CT_Ar08	06° 10' 42.53" S	134° 18' 04.30" E	KY069337	KY069456	KY069553	KY069656	KY069657	KY069726
		CT_Ar09	06° 10' 42.53" S	134° 18' 04.30" E	KY069338	KY069457	KY069554	KY069658	KY069659	KY069727
		CT_Ar10	06° 10' 42.53" S	134° 18' 04.30" E	KY069339	KY069458	KY069555	KY069660	KY069661	KY069728
		CT_Ar11	06° 10' 42.53" S	134° 18' 04.30" E	KY069340	KY069459	KY069556	KY069662	KY069663	KY069729
		CT_Ar12	06° 10' 42.53" S	134° 18' 04.30" E	KY069341	KY069460	KY069557	KY069664	KY069665	KY069730
		CT_Ar13	06° 10' 42.53" S	134° 18' 04.30" E	KY069342	KY069461	KY069558	KY069666	KY069667	KY069731
		CT_Ar14	06° 10' 42.53" S	134° 18' 04.30" E	KY069343	KY069462	KY069559	KY069668	KY069669	KY069732
CT_Ar15	06° 10' 42.53" S	134° 18' 04.30" E	KY069344	KY069463	KY069560	KY069670	KY069671	KY069733		
Bali		ZMB29101	08° 37' 56.70" S	115° 08' 08.58" E	KY069395	KY069514	KY069611	KY069692	KY069769	
		ZMB29344_1	08° 30' 44.76" S	115° 01' 22.20" E	KY069397	KY069516	KY069613	KY069693	KY069770	
		ZMB29344_2	08° 30' 44.76" S	115° 01' 22.20" E	KY069403	KY069522	KY069619	KY069694	KY069771	
		ZMB29344_3	08° 30' 44.76" S	115° 01' 22.20" E	KY069404	KY069523	KY069620	KY069695	KY069772	
		ZMB29344_4	08° 30' 44.76" S	115° 01' 22.20" E	KY069405	KY069524	KY069621	KY069696	KY069773	
		ZMB29344_5	08° 30' 44.76" S	115° 01' 22.20" E	KY069406	KY069525	KY069622	KY069697	KY069774	
		ZMB29344_6	08° 30' 44.76" S	115° 01' 22.20" E	KY069407	KY069526	KY069623	KY069698	KY069775	
		ZMB29344_8	08° 30' 44.76" S	115° 01' 22.20" E	KY069408	KY069527	KY069624	KY069699	KY069776	
		ZMB29344_9	08° 30' 44.76" S	115° 01' 22.20" E	KY069409	KY069528	KY069625	KY069700	KY069777	
		ZMB29344_10	08° 30' 44.76" S	115° 01' 22.20" E	KY069398	KY069517	KY069614	KY069699	KY069778	
		ZMB29344_11	08° 30' 44.76" S	115° 01' 22.20" E	KY069399	KY069518	KY069615	KY069701	KY069779	
		ZMB29344_12	08° 30' 44.76" S	115° 01' 22.20" E	KY069400	KY069519	KY069616	KY069702	KY069780	
		ZMB29344_13	08° 30' 44.76" S	115° 01' 22.20" E	KY069401	KY069520	KY069617	KY069703	KY069781	
ZMB29344_14	08° 30' 44.76" S	115° 01' 22.20" E	KY069402	KY069521	KY069618	KY069704	KY069782			
ZMBuntd1270**	08° 21' 00.00" S	114° 37' 60.00" E	KY069438							

Halmahera	ZMB29416_1	00° 53' 35.22" N	127° 37' 17.22" E	KY069410	KY069529	KY069626	KY069715/KY069716	KY069776
	ZMB29416_2	00° 53' 35.22" N	127° 37' 17.22" E	KY069411	KY069530	KY069627	KY069706	KY069777
	ZMB29416_3	00° 53' 35.22" N	127° 37' 17.22" E	KY069412	KY069531	KY069628	KY069707	
	ZMB29416_4	00° 53' 35.22" N	127° 37' 17.22" E	KY069413	KY069532	KY069629	KY069708	KY069778
	ZMB29416_7	00° 53' 35.22" N	127° 37' 17.22" E	KY069414	KY069533	KY069630	KY069709	KY069779
	ZMB29416_8	00° 53' 35.22" N	127° 37' 17.22" E	KY069415	KY069534	KY069631	KY069710	
	ZMB29011_1	03° 40' 30.06" S	122° 13' 41.88" E	KY069383	KY069502	KY069599	KY069681	KY069756
	ZMB29011_2	03° 40' 30.06" S	122° 13' 41.88" E	KY069384	KY069503	KY069600	KY069682	KY069757
	ZMB29011_4	03° 40' 30.06" S	122° 13' 41.88" E	KY069385	KY069504	KY069601	KY069683	KY069758
	ZMB29011_5	03° 40' 30.06" S	122° 13' 41.88" E	KY069386	KY069505	KY069602	KY069684	KY069759
Sulawesi	ZMB29014_1	04° 23' 17.04" S	122° 35' 33.18" E	KY069387	KY069506	KY069603	KY069685	KY069760
	ZMB29014_5	04° 23' 17.04" S	122° 35' 33.18" E	KY069388	KY069507	KY069604	KY069686	KY069761
	ZMB29014_7	04° 23' 17.04" S	122° 35' 33.18" E	KY069389	KY069508	KY069605	KY069687	KY069762
	ZMB29019_1	03° 40' 59.70" S	122° 15' 17.88" E	KY069390	KY069509	KY069606	KY069688	KY069763
	ZMB29019_2	03° 40' 59.70" S	122° 15' 17.88" E	KY069391	KY069510	KY069607	KY069689	KY069764
	ZMB29091_1	02° 56' 03.00" S	121° 06' 57.60" E	KY069392	KY069511	KY069608	KY069690	KY069765
	ZMB29091_2	02° 56' 03.00" S	121° 06' 57.60" E	KY069393	KY069512	KY069609	KY069691	KY069766
	ZMB29092	05° 00' 57.60" S	119° 40' 55.20" E	KY069394	KY069513	KY069610	KY069711/KY069712	KY069767
	ZMB29206_2	08° 30' 44.76" S	121° 04' 45.00" E	KY069396	KY069515	KY069612	KY069713/KY069714	KY069768
	ZMBBund1273**	04° 04' 37.80" S	121° 40' 04.20" E	KY069439				
Taliabu	DW0850a			KY069351	KY069470	KY069567	KY069658	KY069732
	DW0850b			KY069352	KY069471	KY069568	KY069659	KY069733
	DW0850c			KY069353	KY069472	KY069569	KY069660	KY069734
	DW0850d			KY069354	KY069473	KY069570	KY069661	KY069735
	DW0850e			KY069355	KY069474	KY069571	KY069662	KY069736
	DW0850f			KY069356	KY069475	KY069572	KY069663	KY069737
	DW0850g			KY069357	KY069476	KY069573	KY069664	KY069738
	DW0850h			KY069358	KY069477	KY069574	KY069665	KY069739
	DW0850i			KY069359	KY069478	KY069575	KY069666	KY069740
	DW0850j			KY069360	KY069479	KY069576	KY069667	KY069741
West Papua	DW1607a	00° 12' 06.10" S	130° 48' 48.60" E	KY069361	KY069480	KY069577	KY069668	KY069742
	DW1607b	00° 12' 06.10" S	130° 48' 48.60" E	KY069362	KY069481	KY069578	KY069669	
	DW1607d	00° 12' 06.10" S	130° 48' 48.60" E	KY069363	KY069482	KY069579	KY069670	KY069743

DW1607f	00° 12' 06.10" S	130° 48' 48.60" E	KY069364	KY069483	KY069580	KY069671
DW1607g	00° 12' 06.10" S	130° 48' 48.60" E	KY069365	KY069484	KY069581	KY069672
DW1607h	00° 12' 06.10" S	130° 48' 48.60" E	KY069366	KY069485	KY069582	KY069673
DW1607i	00° 12' 06.10" S	130° 48' 48.60" E	KY069367	KY069486	KY069583	KY069674
DW1607j	00° 12' 06.10" S	130° 48' 48.60" E	KY069368	KY069487	KY069584	KY069675
DW0707			KY069350	KY069469	KY069566	KY069657
ZMB29489_2	06° 24' 07.52" N	099° 49' 10.95" E	KY069416	KY069535	KY069632	KY069780
ZMB29489_3	06° 24' 07.52" N	099° 49' 10.95" E	KY069417	KY069536	KY069633	KY069781
ZMB29489_4	06° 24' 07.52" N	099° 49' 10.95" E	KY069418	KY069537	KY069634	KY069782
ZMB29489_5	06° 24' 07.52" N	099° 49' 10.95" E	KY069419	KY069538	KY069635	KY069783
ZMB29489_6	06° 24' 07.52" N	099° 49' 10.95" E	KY069420	KY069539	KY069636	KY069784
ZMB29489_7	06° 24' 07.52" N	099° 49' 10.95" E	KY069421	KY069540	KY069637	KY069785
ZMB29489_8	06° 24' 07.52" N	099° 49' 10.95" E	KY069422	KY069541	KY069638	KY069786
ZMB29490_1	06° 25' 24.18" N	099° 40' 02.85" E	KY069423	KY069542	KY069639	KY069787
ZMB29490_2	06° 25' 24.18" N	099° 40' 02.85" E	KY069424	KY069543	KY069640	KY069728
ZMB29490_3	06° 25' 24.18" N	099° 40' 02.85" E	KY069425	KY069544	KY069641	KY069729
CT_Sar02	01° 43' 04.71" N	110° 27' 24.43" E	KY069345	KY069464	KY069561	KY069730
CT_Sar03	01° 43' 04.71" N	110° 27' 24.43" E	KY069346	KY069465	KY069562	KY069731
CT_Sar06	01° 43' 04.71" N	110° 27' 24.43" E	KY069347	KY069466	KY069563	KY069748
CT_Sar07	01° 43' 04.71" N	110° 27' 24.43" E	KY069348	KY069467	KY069564	
CT_Sar08	01° 43' 04.71" N	110° 27' 24.43" E	KY069349	KY069468	KY069565	
MEFGL_CT_Kuc01	01° 41' 20.64" N	109° 51' 10.92" E	KY069373	KY069492	KY069589	
ZMB29564	09° 38' 40.14" N	124° 09' 35.52" E	KY069429			
ZMBumid1207**	14° 28' 49.20" N	122° 56' 33.00" E	KY069430			
ZMBumid1235**	14° 47' 50.96" N	120° 19' 03.55" E	KY069431			
ZMBumid1236**	14° 47' 50.96" N	120° 19' 03.55" E	KY069432			
ZMBumid1237**	14° 47' 50.96" N	120° 19' 03.55" E	KY069433			
MEFGL_45_11a_1	14° 47' 18.84" N	120° 16' 12.12" E	KY069369	KY069488	KY069585	KY069676
MEFGL_45_11a_2	14° 47' 18.84" N	120° 16' 12.12" E	KY069370	KY069489	KY069586	KY069677
MEFGL_45_11a_3	14° 47' 18.84" N	120° 16' 12.12" E	KY069371	KY069490	KY069587	KY069678
MEFGL_45_11a_4	14° 47' 18.84" N	120° 16' 12.12" E	KY069372	KY069491	KY069588	KY069679
ZMBumid1258**	12° 37' 07.20" N	121° 24' 06.00" E	KY069435			
ZMBumid839**	10° 48' 52.02" N	119° 30' 42.00" E	KY069441			

Malaysia

Langkawi

Sarawak

Philippines

Bohol

Luzon

Mindoro

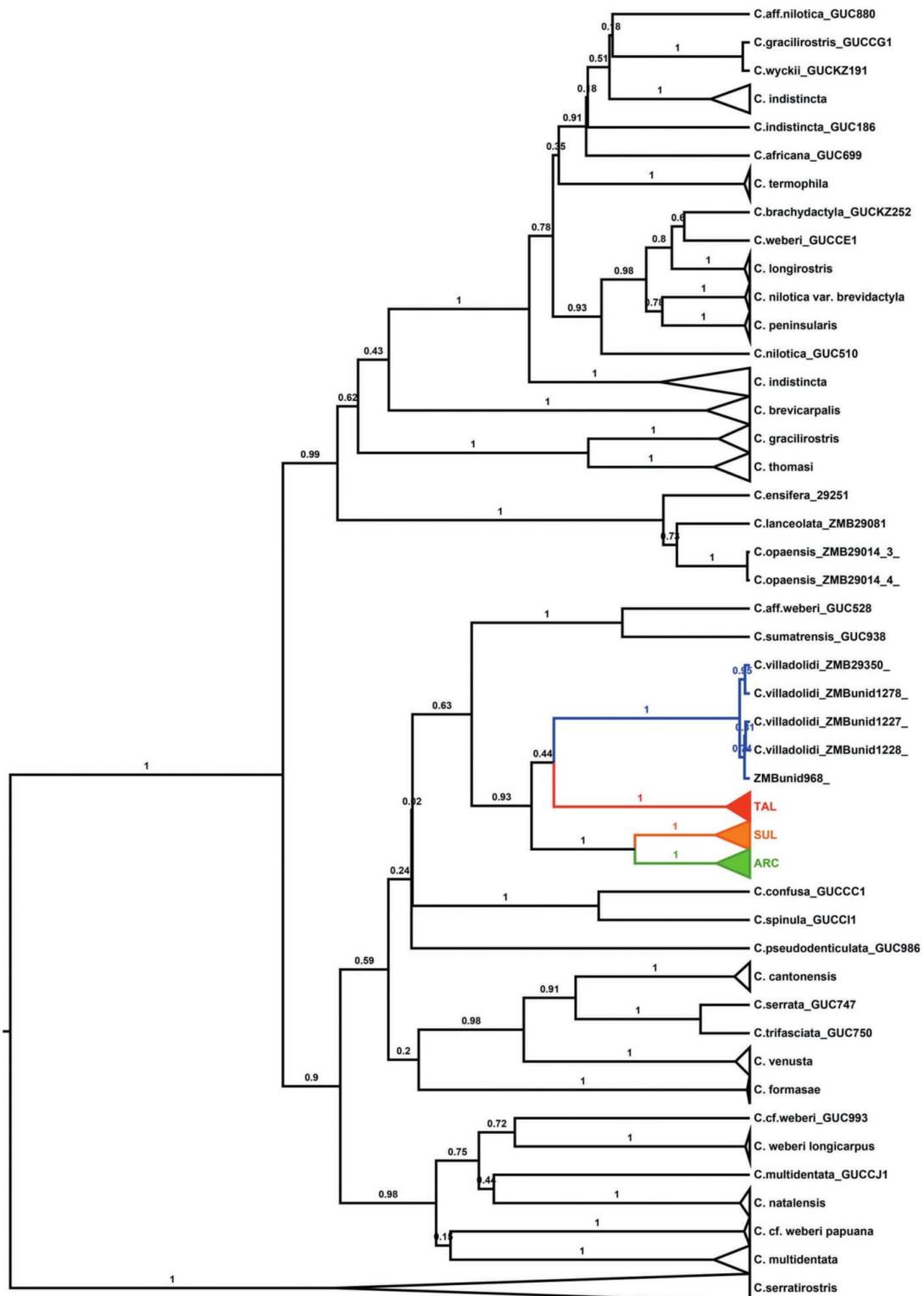
Palawan

	ZMBunid840**	10° 44' 57.00" N	119° 30' 09.00" E	KY069442	
	ZMBunid841**	09° 41' 19.98" N	118° 27' 15.00" E	KY069443	
	ZMBunid842**	09° 46' 36.00" N	118° 39' 55.02" E	KY069444	
	ZMBunid843**	09° 41' 19.98" N	118° 27' 14.28" E	KY069445	
	ZMBunid963**	09° 55' 33.00" N	118° 38' 19.80" E	KY069446	
	ZMBunid1259**	12° 01' 60.00" N	119° 59' 31.20" E	KY069436	
	ZMBunid1267**	12° 15' 25.80" N	119° 59' 46.20" E	KY069437	
Samar	ZMB29547_1	11° 47' 07.02" N	124° 55' 36.00" E	KY069426	
	ZMB29547_2	11° 47' 07.02" N	124° 55' 36.00" E	KY069427	KY069545
	ZMB29548	12° 01' 60.00" N	124° 55' 30.00" E	KY069428	
Taiwan	ZMBunid1283**	22° 12' 00.96" N	120° 52' 42.06" E	KY069440	
Seychelles	MEFGL_CT_Sey01	04° 37' 41.20" S	055° 25' 34.50" E	KY069374	KY069590
	MEFGL_CT_Sey02	04° 37' 41.20" S	055° 25' 34.50" E	KY069375	KY069591
	MEFGL_CT_Sey03	04° 37' 02.70" S	055° 25' 45.00" E	KY069376	KY069592
	MEFGL_CT_Sey04	04° 37' 02.70" S	055° 25' 45.00" E	KY069377	KY069593
	MEFGL_CT_Sey05	04° 37' 02.70" S	055° 25' 45.00" E	KY069378	KY069594
	MEFGL_CT_Sey06	04° 37' 02.70" S	055° 25' 45.00" E	KY069379	KY069595
Mascarene Islands	OUMNH,ZC-2010-22-0001	23° 23' 00.00" S	57° 25' 00.00" E	KY069380	KY069596
	OUMNH,ZC-2010-22-0007	20° 06' 40.00" S	57° 32' 33.00" E	KY069382	KY069598
Rodrigues	OUMNH,ZC-2010-22-0003	19° 42' 00.00" S	63° 24' 30.00" E	KY069381	KY069597
Reunion	ZMBunid1252**	20° 55' 10.50" S	055° 36' 14.88" E	KY069434	
Australia*	GUC571	19° 14' 24.00" S	146° 29' 24.00" E	DQ478561	
	GUCY1	20° 20' 00.00" S	148° 39' 00.00" E	DQ478562	DQ478478.1
Japan*	GUC895	26° 45' 00.00" N	128° 18' 00.00" E	DQ478560	
New Caledonia*	GUC868	20° 20' 02.00" S	164° 18' 51.00" E	DQ478558	
	GUC871	22° 23' 11.00" S	166° 55' 03.00" E	DQ478559	DQ478477.1
Sri Lanka*	FB-2311SL			AY708118	
Vanuatu*	GUC721	14° 24' 00.00" S	145° 16' 12.00" E	DQ478563	
Indonesia	ZMB29350	3° 40' 59.70" S	122° 15' 17.88" E	KY436221	
Philippines	ZMBunid968**	13° 13' 51.00" N	120° 34' 54.00" E	KY436220	
	ZMBunid1227**	12° 27' 51.00" N	121° 02' 07.80" E	KY436222	
	ZMBunid1228**	12° 27' 51.00" N	121° 02' 07.80" E	KY436223	
Taiwan	ZMBunid1278**	21° 59' 18.30" N	120° 49' 57.36" E	KY436224	

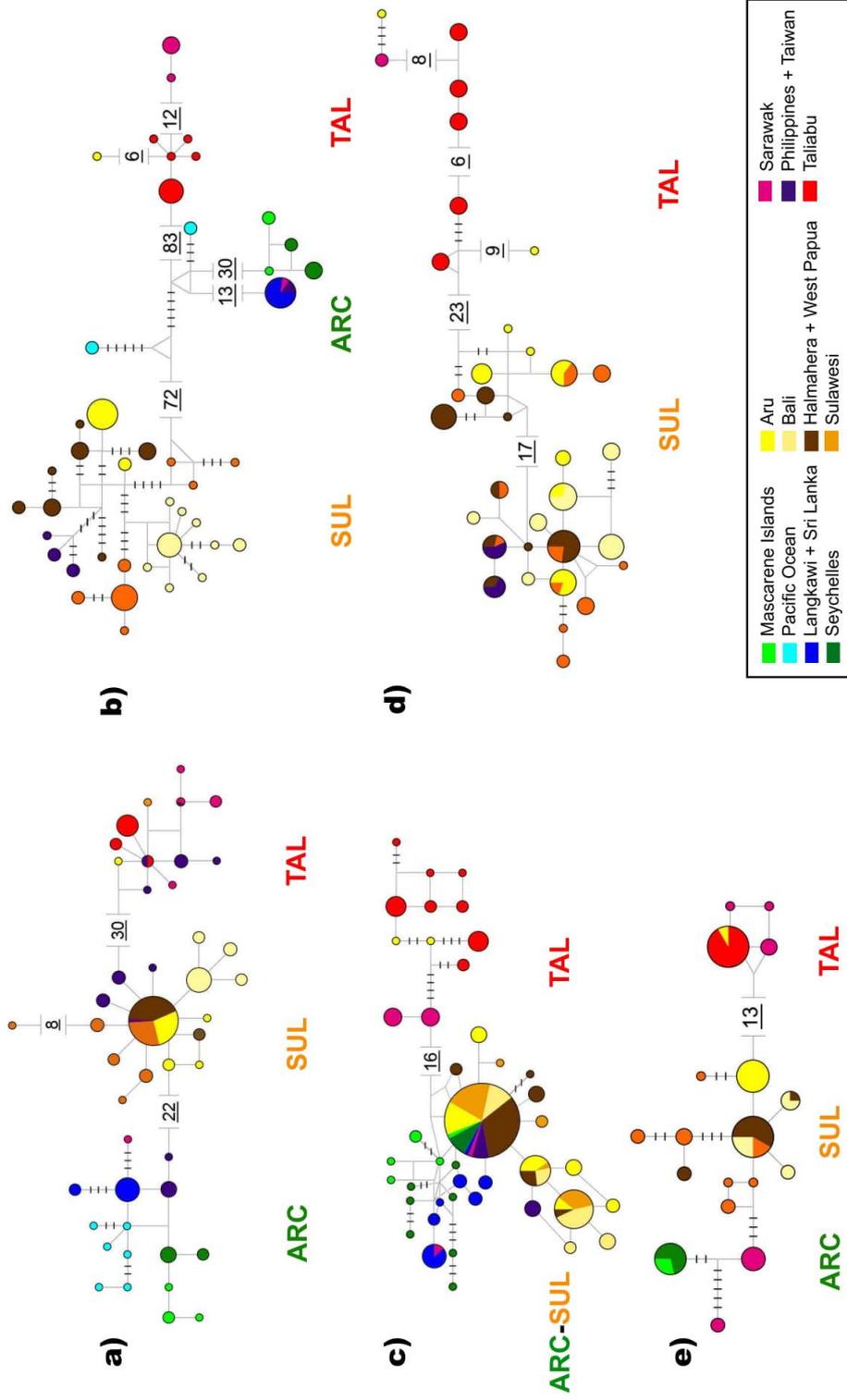
Caridina villadolidi

<i>Caridina cf. opaensis</i>	Indonesia	Sulawesi	ZMB29014_3	04° 23' 17.04" S	122° 35' 33.18" E	KY069328	KY069447	KY069717
			ZMB29014_4	04° 23' 17.04" S	122° 35' 33.18" E	KY069329	KY069448	KY069718

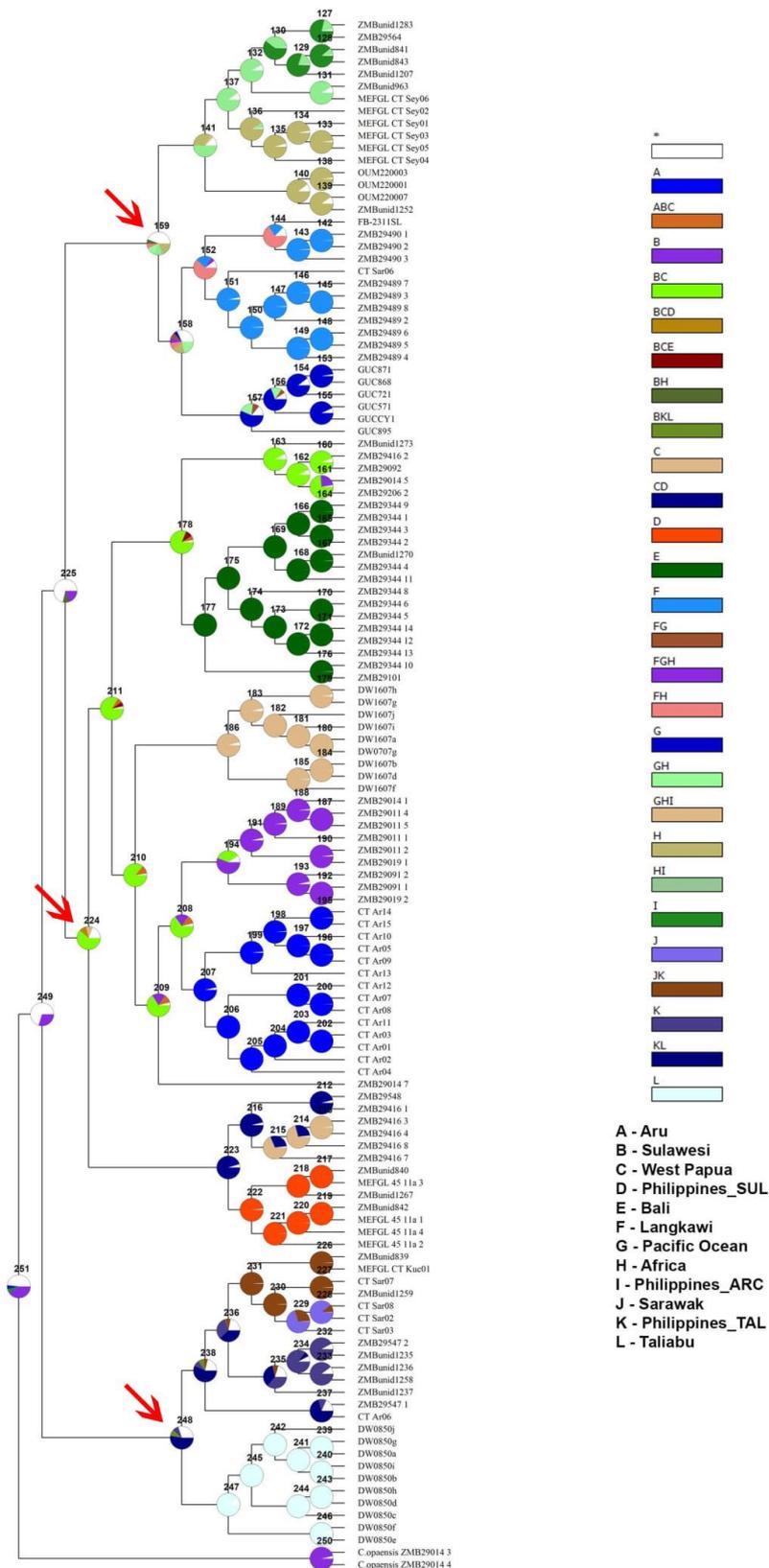
Notes: Countries marked with * had their samples taken from GenBank. With exception of them, these samples were taken from the Museum für Naturkunde Berlin (voucher numbers started in ZMB), Museum Victoria (voucher numbers started in CT_Ar), Oxford University Museum of Natural History (voucher numbers started in OUMNH) and collected samples deposited in the Molecular Ecology and Fisheries Genetics Laboratory in Bangor University (all others). Samples marked with ** had no repository number at the moment of the submission of this paper, but they are deposited in the Museum für Naturkunde Berlin. Samples that have two GenBank accession numbers are heterozygotes and due to differential gaps between alleles, they were deposited twice.



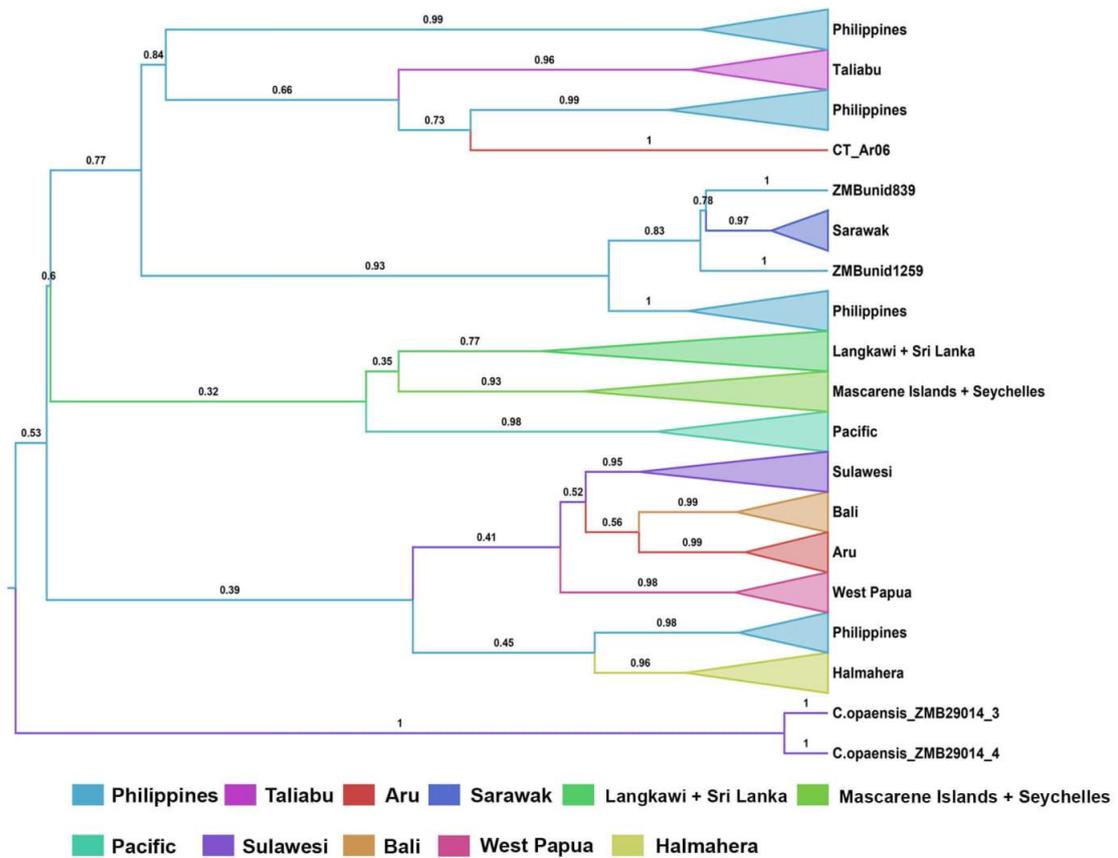
Supplementary Figure S2: Multispecies mitochondrial 16S Bayesian Caridina tree. The values on the branches represent the posterior probability of that clade. *C. typus* complex is coloured: in blue, *C. villadolidi*; in red, TAL; in orange, SUL; and in green, ARC.



Supplementary Figure S3: Haplotype networks for all five marker 16S (a), COI (b), C133 (c), C151 (d), and 28S (e). A single substitution step is represented by a single line connecting two haplotypes. Any number of steps between two and five is represented by the number of dashes on the line connecting two haplotypes. Any distance higher than five steps is represented by the written number of steps. Non-sampled intermediary haplotypes are represented by empty vertices connecting three or more haplotypes. Each colour represents a region according to the legend.



Supplementary Figure S5: RASP's BayArea full result. Each node is represented by a pie chart of relative probabilities between possible ancestral ranges. Each colour in those charts represents a combination of areas according to the legend. The node id is the number above the pie chart. Nodes marked by a red arrow represent the ancestors of each main clade.



Supplementary Figure S6: BEAST discrete phylogeography result with collapsed branches per range. Each colour represents an area, including the colours applied on internal branches. The numbers on the branches represent the posterior probability of a given locality being the ancestor range, given that each terminal has a *pp* of 1.0 for its location.

