UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA COMPARADA

Padrões e processos da diversidade genética em caranguejos (Decapoda, Brachyura): abordagem integrada com atributos biológicos, ecológicos e demográficos.

Patterns and processes of the genetic diversity in crabs (Decapoda, Brachyura): integrative approach with biological, ecological, and demographic traits.

Pedro Augusto da Silva Peres

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para obtenção do título de Doutor em Ciências, obtido no Programa de Pós-Graduação em Biologia Comparada

Ribeirão Preto - SP

2021

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA COMPARADA

Padrões e processos da diversidade genética em caranguejos (Decapoda, Brachyura): abordagem integrada com atributos biológicos, ecológicos e demográficos.

Patterns and processes of the genetic diversity in crabs (Decapoda, Brachyura): integrative approach with biological, ecological, and demographic traits.

Pedro Augusto da Silva Peres

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, como parte das exigências para obtenção do título de Doutor em Ciências, obtido no Programa de Pós-Graduação em Biologia Comparada.

Orientador: Prof. Dr. Fernando Luis Medina Mantelatto

Ribeirão Preto - SP

2021

Peres, Pedro Augusto da Silva
Padrões e processos da diversidade genética em caranguejos (Decapoda, Brachyura): abordagem integrada com atributos biológicos, ecológicos e demográficos., 2021.
iv + 157 p. : il. ; 30 cm

Tese de Doutorado, apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo. Área de concentração: Biologia Comparada.
Orientador: Mantelatto, Fernando Luis Medina.

1. Brachyura 2. Dispersão 3. Filogeografia 4. Genética de populações 5. Latitude 6. mtDNA 7. ddRAD-seq

AGRADECIMENTOS

Ao meu orientador Dr. Fernando Luis Medina Mantelatto pela orientação ao longo de todos os anos do doutorado. Obrigado pela gentileza desde o nosso primeiro contato, por ter me aceitado em seu laboratório e por ter me confiado este projeto com o qual tanto pude aprender. Agradeço imensamente pelos esforços em garantir que eu pudesse ter uma ótima formação, por todas as oportunidades oferecidas, portas abertas e ensinamentos. Acredito que aprendi tanto por suas palavras quanto por suas atitudes e levarei estes aprendizados comigo. Fico orgulhoso de poder fazer parte de seu legado acadêmico.

Aos amigos do Laboratório de Bioecologia e Sistemática de Crustáceos (LBSC) por terem me acolhido quando cheguei, me iniciarem na dinâmica e atividades do laboratório no começo, pela amizade que criamos, por toda a convivência e ajuda em inúmeras situações ao longo desses anos. Sou grato a todos pela amizade, cafés, almoços, reuniões, discussões, bares e rolês que vivemos estes anos.

A Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) por ter financiado a bolsa de doutorado no país (2017/12376-6) e no exterior (BEPE 2018/21127-2), mas também pelo financiamento aos projetos do LBSC concedidos ao Prof. Mantelatto (Temático Biota 2010/50188-8 and Biota INTERCRUSTA 2018/13685-5; Coleções Científicas 2009/54931-0; PROTAX 2016/50376-5), a Coordenação de Aperfeiçoamento de Pessoal—CAPES—Código de Financiamento 001 (Ciências do Mar II—2005/2014–23038.004308/201414) pelo financiamento nos primeiros meses do doutorado e ao LBSC, e também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq (301359/2007-5; 473050/2007-2; 302748/2010-5; 471011/2011-8; PQ 302253/2019-0).

A todos que coletaram e colaboram na manutenção frequente das imprescindíveis coleções cientificas carcinológicas as quais fiz uso durante o projeto, seja por visitas ou empréstimos: Coleção de Crustáceos do Departamento de Biologia (CCDB) da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP/USP), Museu de Zoologia da Universidade de São Paulo (MZUPS), Museu de Oceanografia prof. Petrônio Alves Coelho (MOUFPE) da Universidade Federal de Pernambuco (UFPE), Museu Paraense Emilio Goeldi (MPEG), Florida Museum of Natural History (FLMNH) da University of Florida.

A Dr. Heather Bracken-Grissom e aos amigos do CRUSTOMICS (Crustacean Genomics and Systematics Lab) por terem me recebido e tanto me ensinado na Florida International University e terem feito meu período sanduiche ter sido uma experiência maravilhosa.

A todos os alunos, funcionários e professores das duas universidades públicas que me acolheram e me proveram formação acadêmica e cientifica: Universidade de São Paulo (USP) e Universidade Estadual de Campinas (UNICAMP). Agradeço o empenho na luta em manter ensino e pesquisa pública de qualidade, e por representarem essa grande força de transformação. Um agradecimento especial a todos os professores que pude ter o privilégio de ser aluno na graduação, mestrado e doutorado. “Se eu vi mais longe, foi por estar sobre ombros de gigantes.”

A todos meus amigos que fiz por Bauru, Campinas e Ribeirão Preto, da área acadêmica ou não, por serem minha segunda família, fonte de alegria, inspiração e aprendizado.

A minha família, em especial meus pais Belmiro e Viviani, pelo amor incondicional, por todo apoio, por terem sonhado junto comigo e nunca terem medido esforços para garantir o que eu pudesse trilhar minha própria jornada.

Ao Hudson Ferreira, meu marido, por sempre ter estado por perto, por ter celebrado comigo as conquistas e ter me acolhido nas quedas, pelo que vivemos e vamos viver, e pelo bem que me faz sem perceber.

TABLE OF CONTENTS

RESUMO

Padrões e processos relacionados a diversidade genética (DG) ainda são cercados de questões ecológicas e evolutivas não resolvidas. Entre elas está a que se refere aos fatores que regulam a variação da DG dentro e entre as diferentes espécies. Tal tema ainda carece de maior quantidade de dados empíricos, principalmente em relação ao megadiverso ambiente marinho. Nesta tese, utilizamos espécies de caranguejos (Eubrachyura) como modelos para avaliar como atributos biológicos, ecológicos e demográficos podem estar relacionadas à DG. Esta tese é composta por quatro capítulos (um artigo publicado e três em preparação) que visam responder diferentes questões sobre o tema. No capítulo 1, testamos a hipótese de que características biológicas podem ser mais relevantes que o potencial de dispersão de organismos para explicar padrões de distribuição da DG comparando duas espécies de caranguejos filogeneticamente próximas (*Callinectes ornatus* e *C. danae*). Como resultado, mostramos que a tolerância a salinidade pode estar relacionada à estrutura genética, padrões filogeográficos e história demográfica de ambas as espécies. A partir de perguntas levantadas com os resultados anteriores, no capítulo 2 investigamos os efeitos da foz do Amazonas-Orinoco (barreira fisiológica) sobre *C. ornatus,* utilizando dados obtidos a partir de sequenciamento de nova geração (ddRAD-seq) combinados com mtDNA. Como resultados, mostramos um cenário de especiação com fluxo gênico em que grupos separados pela foz estão bastante diferenciados, mas ainda com fluxo gênico existente entre eles. No capítulo 3, investigamos os padrões espaciais da DG utilizando dados de mtDNA de 14 espécies de caranguejos ao longo de toda sua distribuição no Atlântico Ocidental e testamos a hipótese de gradiente latitudinal da DG. Encontramos que a diversidade genética interespecífica é maior em baixas latitudes, mas que os padrões intraespecíficos variam de acordo com a espécie. Por fim, no capítulo 4, investigamos a relação entre características biológicas e

demográficas (8 atributos) com a DG. Utilizamos todos os dados de mtDNA disponíveis para Eubrachyura em bancos de dados públicos junto com dados gerados durante este projeto (150 espécies), e realizamos busca padronizada das características biológicas. De maneira geral, encontramos que o tamanho populacional efetivo é o fator que mais explica a variação da DG em caranguejos, mas alguns outros atributos parecem ter importância. Portanto, temos a convicção de que esta tese abre novos horizontes a partir de propostas inéditas e relevantes sobre padrões e processos da DG utilizando caranguejos como modelo, mas também podem se estender a outros táxons, especialmente marinhos.

PALAVRAS-CHAVE: Brachyura; Dispersão; Filogeografia; Genética de populações; Latitude; mtDNA; ddRAD-seq.

ABSTRACT

The patterns and processes of the genetic diversity (GD) are still surrounded by unresolved questions in the fields of ecology and evolution. One of the questions is what drives GD at the intra- and interspecific level. This topic needs further empirical data, especially regarding the megadiverse marine environment. In this dissertation, we used crabs (Eubrachyura) as models to test how biological, ecological, and demographic traits are associated to GD. This dissertation is composed of four chapter (one published paper, and three papers in preparation) that asked different questions within this framework. In chapter 1, we compared two phylogenetically close species (*Callinectes ornatus* and *C. danae*) to test the hypothesis that biological traits are more important than the dispersal potential explaining GD spatial patterns. We show that salinity tolerance can explain genetic structure, phylogeographic patterns, and the demographic history of both species. Considering the questions that remained unanswered, in chapter 2 we investigated the effects of the Amazon-Orinoco plume (physiological barrier) on *C. ornatus* employing a next-generation sequencing approach (ddRAD-seq) alongside with mtDNA. We show a speciation with gene flow scenario in that groups separated by the plume are differentiated, but we still find gene flow between them. In chapter 3, we investigated the GD spatial patterns using mtDNA of 14 crabs from the Western Atlantic and tested the latitudinal gradient hypothesis. Our results show that interspecific GD is higher at lower latitudes, but intraspecific patterns vary across species. Finally, in chapter 4, we investigated the association among biological and demographic traits (8 variables) with GD. We analyzed the most comprehensive dataset to date of Eubrachyura mtDNA available in public databases and generated during this dissertation (150 species) and did standardized searches in the literature looking for the biological traits. Overall, our results show that the historical effective population size is the primary factor determining crabs GD variation, but other

traits might also contribute to this variation. Therefore, we are sure this dissertation opens new venues by providing novel approaches on GD patterns and processes using crabs as models, but our results might also be extended to other taxa, especially marine ones.

KEY WORDS: Brachyura; ddRAD-seq; Dispersion; Latitude; mtDNA; Phylogeography; Population genetics.

GENERAL INTRODUCTION

The genetic diversity (GD) is a main component of biodiversity. GD can be defined as "the variation in a DNA sequence between distinct individuals (or chromosomes) of a give species (or population)" (Ellegren & Galtier, 2016). GD is the trait responsible for species' adaptive potential when facing environmental changes (Eizaguirre & Baltazar-Soares, 2014) and for preventing endogamic depression due to the accumulation of deleterious mutations (O'Grady et al., 2006). GD also has effects above the population or species level, tending to be correlated with community composition and ecosystem functioning (e.g., productivity, decomposition) in terrestrial and marine environments (Hughes et al., 2008; Whitlock, 2014; Jormalainen et al., 2017). Indeed, it is recognized by the Convention on Biological Diversity (CBD, https://www.cbd.int/) as one of the levels of biodiversity to be conserved and sustainably used.

Since the advent of molecular biology, investigations on GD questions have witnessed significant gains due to DNA sequencing technologies improvements (Schlötterer, 2004; Mardis, 2013). We have accumulated thousands of sequences that are available in public databases (Porter & Hajibabaei, 2018). Additionally, the amount of data and information content are increasing exponentially over the years, especially during the past two decades, and the new platforms are faster and can be relatively cheaper comparing with traditional Sanger method (McCormack et al., 2013). Therefore, we are living a turning point in ecological and evolutionary studies that employ molecular markers. Datasets are getting bigger because Sanger sequences data can be generated and combined with available sequences, and because we can generate hundreds of thousands of markers using next-generation sequencing (NGS) technologies (Rowe et al., 2011). However, the understanding what drives species GD is still one of the open questions in ecology and evolution (Leffler et al., 2012; Ellegren & Galtier, 2016). Some of the intriguing questions are what the GD

spatial patterns are and what the determinants of GD are, so as the processes generating these patterns.

GD spatial patterns can be investigated at the intraspecific level. Traditionally, marine systems have been considered more "open" due to the transport of particles by currents and the wide distribution of species, resulting in higher import and export rates among regions (Palumbi, 1992). Many marine species show a larval phase during their life cycle. The larvae are released by adults in the water and have the potential to be transported by ocean currents and reach distant locations (Hedgecock, 1986).

Long pelagic larval duration (PLD) has been considered a trait that ensures the genetic homogeneity between interconnected populations (Ayre et al., 1997; Shanks, 2009). In general, there is a positive association between the pelagic larval duration (PLD) and the distance traveled by the larvae (Shanks, 2009). Consequently, longer PLD would result in a lack of population differentiation, but this is not always true.

However, this trait may not be the solely predictor of genetic structure, as shown by several results that highlight the importance of other natural history traits (Baus et al., 2005; Kenchington et al., 2006; Ayre et al., 2009; Weersing & Toonen, 2009; Teske et al., 2011). Larval retention caused by local currents, larval behavior, and larval development rate variation depending on the region, or adult's ecology is among the factors explaining why a long PLD not always result in connectivity among distant populations (White et al., 2010; Butler et al., 2011 Hedgecock & Pudovkin, 2011; Álvarez-Noriega et al., 2020; Timm et al., 2020). Thus, even phylogenetically close and co-distributed species with similar PLD can also show contrasting genetic patterns (Eble et al., 2009). Another common assumption in marine species is that organisms lack phylogeographical patterns due to absence of physical barriers to gene flow (Palumbi, 1992).

Again, this assumption has been contradicted by the presence of soft barriers (Briggs & Bowen, 2013) and evidence of cryptic speciation in marine habitats (Palumbi, 1992; Bohem et al., 2013; Negri et al., 2014). The recent implementation of NGS approaches also have revealed fine-scale and adaptive structure in species showing long PLD (Saenz-Agudelo et al., 2015; Benestan et al., 2015; Xuereb et al., 2018; Teske et al., 2019). Therefore, marine species represent potential model organisms to investigate such questions.

GD spatial patterns can also be investigated at the interspecific level. Species richness tends to be higher towards the tropics on a global scale, a spatial pattern known as the Latitudinal Diversity Gradient (LDG), one of the most intriguing and well documented biological patterns (Hawkins, 2001; Hillebrand, 2004a; Kinlock et al., 2017). Although there is no consensus on the processes determining this pattern, most explanations fall within historical, biogeographical, and/or ecological processes (Mittelbach et al., 2007; Lawrance & Fraser, 2020). Recent studies addressing if the LDG extends to the GD found support for a broad-scale Latitudinal Genetic Diversity Gradient (LGDG) (Miraldo et al., 2016; Gratton et al., 2017; Schlutter & Pennel, 2017). As the LDG, the LGDG shows a spatial pattern of higher GD at lower latitudes and decreasing GD towards higher latitudes. However, the LGDG remains poorly explored, and there are unresolved questions about whether this is a general trend across different taxa and the relationship between the intra- and interspecific levels.

Most of the studies on the GD spatial distribution deal with terrestrial vertebrates or invertebrates from temperate regions (Eckert et al., 2008). We still lack information about the patterns and processes of GD distribution across marine species (but see Manel et al., 2020 for a discussion on fishes), mainly tropical marine invertebrates (but see Liggins et al., 2014 for a discussion on echinoderms). The LDG is controversial when considering the marine environment.

We find varying support in favor of the LDG depending on the taxa and region (Hillebrand, 2004b; Titternsor et al., 2010), but there are also cases showing the opposite response – a latitudinal inverse gradient (Rivadeneira et al., 2011). Furthermore, there is an indication of a bimodal latitudinal species gradient in marine environments, which shows higher diversity at intermediate latitudes (Chaudhary et al., 2016; Chaudhary et al., 2017), not following LDG. If GD follows the species richness gradient, we may find different marine environment patterns than those described for terrestrial environments. Again, marine species represent potential candidates to explore these questions due to their underrepresentation within this framework.

Habitat has also been shown as an important factor for GD at the interspecific level. Upland Amazonian bird species show higher GD than floodplain species (Harvey et al., 2017), terrestrial birds show higher GD than aquatic birds (Eo et al., 2011), marine fishes show higher GD than freshwater species (DeWoody & Avise, 2000; Martinez et al., 2018), and shallow decapod species show higher GD than deep-sea species (García-Merchán et al., 2012). Similar habitats may undergo the same geological and abiotic changes leading to similar demographic responses influencing the GD of the habitat-associated fauna in terrestrial and marine environments (Marko et al., 2010; Gehara et al., 2017). Some habitat types are more connected through species dispersal resulting in patterns like canopy bird species being less genetically differentiated than understory species (Burney et al., 2009) and less genetic differentiation explained by depth in marine animals (Etter et al., 2005; García-Merchán et al., 2012; Selkoe et al., 2014). However, we still have to consider idiosyncratic responses due to species from the same habitat do not always share the same traits (e.g., physiology, dispersal potential) (Buckley, 2009; Papadopoulou & Knowles, 2016).

As we mentioned, another question is what determines the GD variation across species (Leffler et al., 2012). The neutral theory of molecular evolution predicts that GD is proportional to the

effective population size (N_e) at neutral sites because of the mutation/drift equilibrium. Hence, the bigger the N_e, the bigger the GD. This relationship has been tested empirically using approximations of N_e and different markers throughout the years (Soule, 1976; Frankham et al., 1996; Montgomery et al., 2000; Romiguier et al., 2014; Mackintosh et al., 2019). The population size-GD relationship usually indicates that groups known to have larger population sizes show higher nucleotide diversity (e.g., insects > mammals) (Frankham et al., 1996; Leffler et al., 2012). Some inconsistencies are found when mitochondrial DNA (mtDNA) is the marker of choice investigated. There is support for the evidence of mtDNA GD and population size being proportional (Mulligan et al., 2006; Nabholz et al., 2008a; Piganeau & Eyre-Walker, 2009), so as for no relationship at all (Bazin et al., 2006), leaving this question demanding further investigation.

Recently, studies have pointed out that life-history traits might be more important than the N_e to determine GD variation (e.g., Romiguier et al., 2014; Kort et al., 2021). The authors argue that traits like propagule size, body size, fecundity, and others are potentially related to N_e, offering a more accurate way to investigate the intensity of demographic changes over time that resulted in the GD found nowadays. However, the current knowledge on animals GD is still based on a restricted group of animals, especially vertebrates, and we might be missing new trends due to the lack of investigation on neglected taxa. For instance, bony fishes show a negative relationship between GD and maximum size, egg diameter, and length at maturity as expected (Mitton & Lewis Jr., 1989), but some butterfly families, though showing a negative correlation between GD and size, show no relationship between GD and egg size, larval host plat and current abundance (Mackintosh et al., 2019). Although we may find general trends, taxon-related patterns may emerge.

Considering all that we discussed above, crabs (Brachyura) emerge as a potential group to do such kind of studies. They are one of the most diversified invertebrates and one of the most studied crustaceans (Ng et al., 2008; Davie et al., 2015; Wolfe et al., 2021). Crabs are found from abyssal zones to terrestrial environments occupying most habitats, showing a vast life-history traits variation and dispersal potential (Hines, 1982; Hines, 1986; Ng et al., 2008; Anger et al., 2015; Davie et al., 2015). Hence, they represent interesting model organisms to investigate questions on the GD spatial patterns and the determinants of the GD at many levels.

This dissertation is composed of independent but related four chapters addressing different questions and testing different hypotheses on patterns and processes related to the GD using crabs as model organisms. All chapters are presented in the format of scientific articles. In the first chapter, already published, we chose two co-distributed and closely related swimming crab species (*Callinectes ornatus* and *C. danae*) to test the hypothesis of an ecological trait being stronger at predicting phylogeographic patterns than the dispersal potential. These species are distributed along the tropical western Atlantic and show different salinity tolerance (*C. danae* > *C. ornatus*). We tested the hypothesis that the Amazon-Orinoco plume (the largest freshwater and sediment discharge into the ocean in the world) represents a barrier for *C. ornatus*, but not for *C. danae.* In the second chapter, we decided to go deeper into the Chapter 1 questions. We used a NGS approach (ddRAD-seq) to get thousands of SNPs to investigate neutral and adaptive structure between the regions at North and South of the Amazon-Orinoco plume. Our results helped us to better understand the complex diversification scenario within *C. ornatus* lineages by indicating the processes involved in shaping the patterns we have found. In the third chapter, we tested the hypothesis of LGDG in 14 crab species from the western Atlantic and if the evolutionary speed hypothesis (ESH – Rohde, 1992) is the process behind a possible pattern. As alternative

hypotheses, we tested the central marginal hypothesis (CMH – Eckert et al., 2008), the CMH-LGDG hypothesis (Guo, 2012) or the marine species hypothesis (MH – Liggins et al., 2015) could better explain the GD spatial pattern in crabs. Finally, in the fourth chapter, we compiled the most comprehensive cytochrome c subunit I (COI) dataset to date (150 species, 16992 sequences) and tested the effect of different life-history and demographic variables that potentially influence GD.

Evolutionary Ecology (2020) 34:589–609
https://doi.org/10.1007/s10682-020-10057-x

ORIGINAL PAPER

Salinity tolerance explains the contrasting phylogeographic patterns of two swimming crabs species along the tropical western Atlantic

Pedro A. Peres[1] · **Fernando L. Mantelatto**[1]

Received: 18 January 2020 / Accepted: 4 June 2020 / Published online: 19 June 2020

Abstract
Patterns and processes of species diversification in the oceans are still not fully understood. Traditionally, studies have been using the pelagic larval duration (PLD) to explain the genetic structure and phylogeographic history of marine taxa. However, this trait has given inconsistent results, especially when there is a physiological barrier. Phylogeographic studies comparing species that have similar PLD but differ in other important traits can indicate which ones drive intraspecific evolution. To test our hypothesis, we selected two species with similar distribution and PLD and different salinity tolerance to explore the role of Amazon-Orinoco plume (the biggest freshwater discharge into the ocean worldwide) in the diversification of western Atlantic species. We amplified mtDNA markers (COI and 16S rRNA) of *Callinectes ornatus* (less tolerant to low salinity) and *C. danae* (tolerant to low salinity) from both sides of the Amazon-Orinoco plume (four biogeographical provinces). Then, we performed genetic structure, historical demography, divergence time, and biogeographic modelling analyses. Our results show contrasting phylogeographic and demographic patterns that can be explained by salinity tolerance. The Amazon-Orinoco plume represents a barrier for *C. ornatus*, which has two evolutionary units (ESUs). The plume is not a barrier for *C. danae*, which has no genetic structure. Furthermore, *C. ornatus* is formed by an ancestral Caribbean group that dispersed to the southwestern Atlantic after the establishment of the Amazon-Orinoco plume. *Callinectes danae* has undergone demographic changes during the Last Glacial Maximum, probably due to the loss of estuarine habitats due to sea level fall, while *C. ornatus* was not affected because it is absent in this type of environment. Therefore, we show that ecological traits of marine taxa, like salinity tolerance, are more reliable predictors of genetic variation than the usually used larval dispersal potential.

Keywords Biogeography · *Callinectes* · Comparative phylogeography · Dispersion · mtDNA · Portunidae

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10682-020-10057-x) contains supplementary material, which is available to authorized users.

✉ Pedro A. Peres
pedro.peres27@gmail.com

Extended author information available on the last page of the article

Introduction

Despite the methodological and theoretical advances in phylogeography since its foundation (Avise et al. 1987; Hickerson et al. 2010), researchers still lack a deep understanding of marine phylogeography (compared to the terrestrial environment), especially regarding marine invertebrates (Beheregaray 2008). Traditionally, long pelagic larval duration (PLD)—a common biological feature of marine species—has been considered a trait that ensures the genetic homogeneity between interconnected populations (Ayre et al. 1997; Shanks 2009). However, this trait may not be the solely predictor of genetic structure, as shown by several results that highlight the importance of other natural history traits (Baus et al. 2005; Kenchington et al. 2006; Ayre et al. 2009; Weersing and Toonen 2009; Teske et al. 2011). Another common assumption in marine phylogeography is that organisms lack phylogeographical patterns due to absence of physical barriers to gene flow (Palumbi 1992). Again, this assumption has been contradicted by the presence of soft barriers (Briggs and Bowen 2013) and evidences of cryptic speciation in marine habitats (Palumbi 1992; Bohem et al. 2013; Negri et al. 2014). Thus, marine invertebrates are exciting models for phylogeographical studies since there are many knowledge gaps to be explored (Palero et al. 2008; Sotelo et al. 2009; Kelly and Palumbi 2010), especially in highly diverse tropical areas (Beheregaray 2008).

The tropical Atlantic is a puzzling biogeographic realm (Joyeux et al. 2001) that extends from the African to the American coast, and comprises seven provinces and potential geographical and physiological barriers (Spalding et al. 2007). For instance, the mid-Atlantic barrier and the Orinoco-Amazon plume are two of the biogeographical barriers found in the tropical Atlantic (Floeter et al. 2008). However, the influence of these barriers on marine taxa depends on the species (Joyeux et al. 2001), which hampers the detection of common processes driving the species distribution and population connectivity. Comparative phylogeography studies, in this sense, can shed light on how historical events or barriers are the causal factors of shared phylogeographical patterns among co-distributed species (Avise 2000; Arbogast and Kenagi 2001). These studies usually search for congruent patterns among taxa and, when necessary, discuss contrasting patterns a posteriori (Burton 1998; Dawson 2001; Hoareau et al. 2013). Due to empirical advances and insights on how historical events influence intraspecific lineages, some authors have argued in favor of a different approach: instead of congruent patterns, one should search for discordant patterns (Buckley 2009; Papadopoulou and Knowles 2016). That means, formulating a priori hypotheses based on species-specific traits, stressing this intrinsic information as causal factors of the organization of genetic lineages or genetic variation. Marine phylogeography hypotheses, specifically, are commonly based on the dispersal potential of larval phases, which is a controversial trait (Weersing and Toonen 2009). Thus, trait-based hypotheses may explain puzzling patterns found in the marine realm and help to predict how environmental changes may affect genetic variation in the future.

Considering the background exposed above, the study design should take into account co-distributed species with already available ecological trait information (Bell et al. 2017), and that also vary in a target trait. The commercial swimming crab species *Callinectes ornatus* Ordway, 1863 and *Callinectes danae* Smith, 1869, for instance, represent an interesting pair of coexisting species that can be used to explore trait-based hypotheses. Both species occur in the shallow waters of the western tropical Atlantic, from Florida to southern Brazil (Mantelatto and Fransozo 2000; Chacur and Negreiros-Fransozo 2001). While

their complete larval development is unknown, we may infer they have up to nine larval phases and a long PLD, based on their larval morphology, and analogy with other congeners and members of Portunidae (Costlow and Bookhout 1959; Bookhout and Costlow 1977; Mantelatto et al. 2014). Nevertheless, these species differ in an important ecological trait of marine organisms: salinity tolerance (Negreiros-Fransozo and Fransozo 1995). Both species occurs from intertidal zone up to 75 m depth, but *Callinectes ornatus* lives in habitats of moderate to high salinities, whereas *C. danae* occurs frequently in estuarine areas (Norse 1978; Mantelatto and Fransozo 2000). Egg hatching of *C. ornatus* occurs in marine saline open areas, while in *C. danae* it occurs in less saline areas, adjacent to the coast (Mantelatto 2000; Keunecke et al. 2012; Andrade et al. 2014). Furthermore, the sodium pump activity, which maintains cell osmotic and ionic equilibria, is threefold lower in *C. ornatus* than in *C. danae* (Masui et al. 2002; Garçon et al. 2007), reinforcing the idea that these ecological differences are driven by physiological differences.

The ecological differences between *C. ornatus* and *C. danae* make this species pair a good model to answer the question of whether their physiological constraints result in contrasting phylogeographical patterns along their distribution. There is a putative physiological marine barrier in the middle of their distribution (north-eastern portion of South America): The Amazon-Orinoco plume. The Amazon-Orinoco plume constitutes the biggest freshwater and sediment discharge into the ocean worldwide (Curtin 1986), which alters local ocean conditions. The decrease in salinity created by this discharge affects up to 30 m in depth and as far as 400 km off coast (Pailler et al. 1999). The establishment of this plume, around 7 Mya, was a consequence of the inversion of the Amazon freshwater system caused by the Andes uplifting ~ 10 Mya (Hoorn et al. 2010). It is known that the plume represents a barrier for some species (Nunes et al. 2017; Silva et al. 2018) but also does not affect others (Joyeux et al. 2001). It is important to notice that there is a reef habitat under the plume (Moura et al. 2016), which is an environment where *Callinectes* are hardly find (Williams 1974). Thus, dispersal may only be possible via surface waters for both species, and not by under the plume (Rocha et al. 2002). Therefore, our dataset provides an opportunity to assess the influence of this barrier on congeneric species with known differences in salinity tolerance.

Due to the different salinity tolerance we expected to find different phylogeographic patterns. Traditionally, and due to long PLD, we would expect a lack of phylogeographical patterns in both species and high levels of connectivity. Here, our hypothesis is that *C. ornatus* is formed by two evolutionary significant units (ESUs) because the Amazon-Orinoco plume acts as an effective physiological barrier for this species. ESUs are defined as intraspecific units reproductively isolated from other units, and can be identified as reciprocally monophyletic lineages for mtDNA (Waples 1991; Moritz 1994, but see Crandall et al. 2000; Fraser and Bernatchez 2001 for review). The two ESUs diverged with the establishment of the Amazon-Orinoco and, consequently, underwent independent demographic histories. On the other hand, we expected that the freshwater outflow would not be an effective physiological barrier to *C. danae*, and that this species would be composed of a single ESU whose populations share the same demographic history along the distribution range.

Materials and methods

Sampling

The sampling scheme was designed to cover the entire distribution range of *Callinectes ornatus* and *C. danae*, including individuals from both sides of the Amazon-Orinoco plume and from the four western Atlantic provinces (Fig. 1, Table S1). Museum specimens were

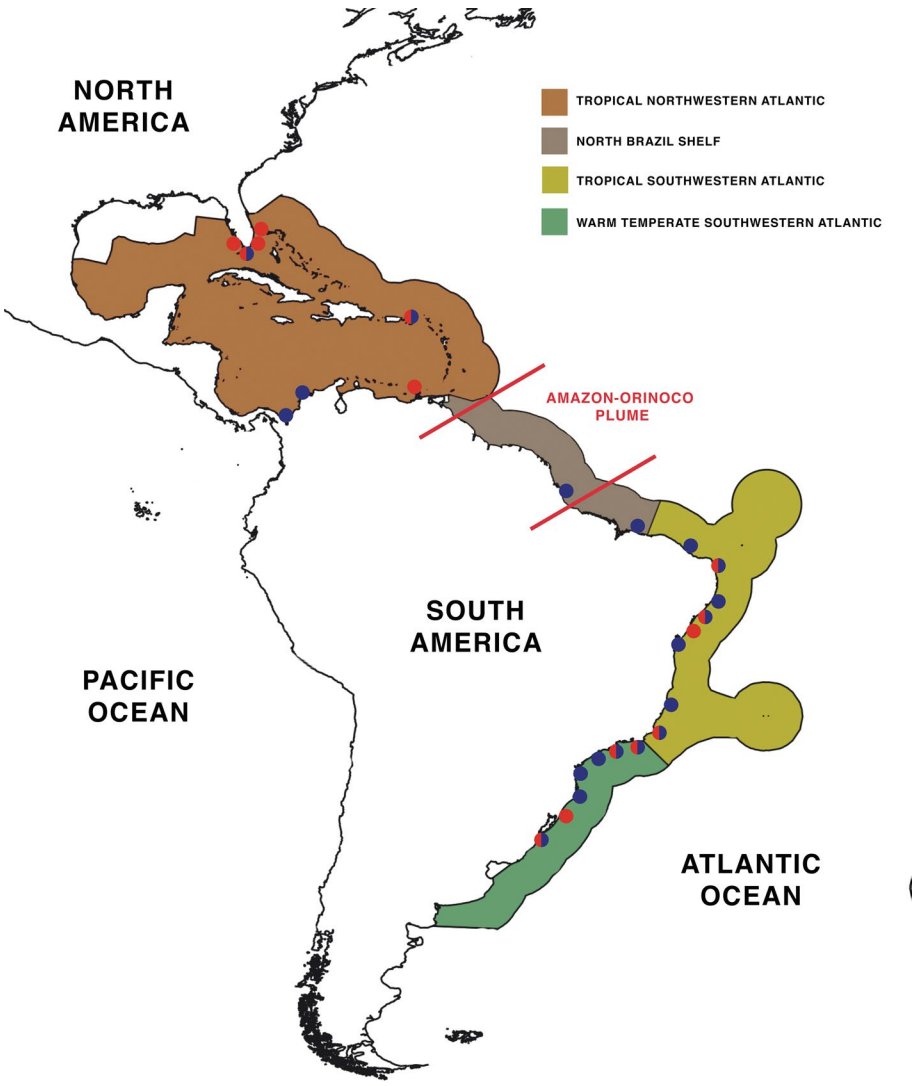


Fig. 1 Map of western Atlantic showing biogeographical provinces (sensu Spalding et al. 2007) where sampling was performed. Each biogeographical province is represented in a different color. Red lines represent Amazon River and Orinoco River mouth, resulting the Amazon-Orinoco Plume. Blue dots represent *Callinectes danae*, and red dots represent *C. ornatus*. Dots half blue half red are sites where both species were sampled

obtained from the Crustacean Collection of the Department of Biology (CCDB), Faculty of Philosophy, Sciences and Letters at Ribeirão Preto (FFCLRP) of the University of São Paulo, and from the Invertebrate Zoology Collection, Florida Museum of Natural History (FLMNH) of the University of Florida.

DNA extraction, amplification, and sequencing

Genomic DNA was extracted from muscle tissues using the salt extraction method (Miller et al. 1988), with the modifications indicated in Mantelatto et al. (2006).

Two mitochondrial (mtDNA) molecular markers were used in our analyses. Sequences of *C. danae* were generated by our research group during previous studies (Robles et al. 2007; Peres et al. 2020) and were used here in combination with new generated sequences. Sequences of *C. ornatus* were all generated during this study. The cytochrome oxidase I (COI) was amplified with the primer pair COH6/COL6b (Schubart and Huber 2006), and the 16S rRNA (16S) was obtained with 16SL2/16S1472 (Schubart et al. 2000; Crandall and Fitzpatric 1996). These markers have been commonly used in decapod phylogeographic studies (Negri et al. 2018; Mandai et al. 2018; Buranelli and Mantelatto 2017; Parvulescu et al. 2019; Oliveira et al. 2019). PCR protocols were performed as indicated in Magalhães et al. (2016). The amplicons were purified using the SureClean Plus® kit, and both strands were sequenced with the ABI Big-Dye Terminator Mix (Applied Biosystems®) in an ABI 3730xl DNA Analyzer (Applied Biosystems® automated sequencer).

We performed the quality filtering, trimming, primer removal, and *denovo* assembling of strands using Geneious v11.1.4 (Kearse et al. 2012). We checked for pseudogenes in COI consensus sequences (protein-coding sequence) by translating them and checking for indels and stop codons (Song et al. 2008) and removed those sequences from the analysis. We aligned the sequences using MAFFT v.7 (Katoh and Standley 2013), which resulted in final alignment of ~ 550 base pairs (bp) for 16S rRNA and ~ 600 bp for COI mtDNA. Our final dataset consisted in both genes for all individuals, which was used for all analyses.

Genetic structure analysis, genetic diversity and haplotype network

We investigated the genetic structure of *C. danae* and *C. ornatus* using the COI and 16S concatenated sequence. A Bayesian clustering method was used to assign individuals to panmictic units and unveil latent structure. Each unit was considered an ESUs as they represent an intraspecific reproductively isolated unit. We employed a Bayesian Analysis of Population Structure—BAPS v.6 (Corander et al. 2003, 2004; Corander and Tang 2007). First, we ran an analysis of population mixtures, setting k as 1, 5, 10, 15 and 20. Then, we ran an individual-based admixture analysis using the results of the previous analysis, to estimate discrete groups. We repeated the analysis 100 times per individual. DNASP v.4.10.9 (Rozas and Rozas 1999) was used to obtain summary statistics such as number of haplotypes (n), polymorphic sites (S), haplotypic diversity (*h*), and nucleotide diversity (π), of each biogeographical province and for each panmictic population.

Additionally, the relationship between haplotypes was estimated using a statistical parsimony network through the TCS method (Clement et al. 2000) implemented in PopArt 1.7 (Leigh and Bryant 2015).

Demographic history analysis

Due to the assumptions of coalescent methods, the use of structured populations may lead to wrong results, thus, we tested for changes in population size using only panmictic populations for both species (i.e. based on the BAPS analysis; see the Results section) (Rosenberg and Nordborg 2002). Neutrality tests were performed in DNASP using Tajima's D (Tajima 1989), Fu's Fs (Fu 1997) and R2 statistics (Ramos-Onsins and Rozas 2002) using the concatenated sequences. The significance of each test was calculated with 10,000 coalescent simulations. Additionally, we reconstructed the historical demography of each population over time using the Bayesian Skyline Plot (BSP) (Drummond et al. 2005) implemented in BEAST2 (Bouckaert et al. 2014). We linked trees and the clock model of markers because both are mitochondrial genes. We used a Relaxed log-normal clock and used the mtDNA mutation rate of ~ 2.3% Mya, which is the standard for crustaceans (Knolwton et al. 1993), setting substitution rate as 0.010–0.015. The same mutation rate was set for each panmictic population considered for this analysis (i.e. based on the BAPS analysis results). We used jmodeltest 2.1.10 (Posada 2008) to find the best substitution model for each marker, which were: HKY for *C. danae* COI and 16S; HKY + G (for COI) and HKY + I (for 16S) for *C. ornatus*. Depending on the population we had to use different chain length and logging to achieve convergence. In the case of *C. danae* we performed a run with 50 million generations, with trees sampled every 5000 generations. In the case of *C. ornatus* there were 10 or 50 million generations, with trees sampled every 1000 generations, depending on the population. All runs were performed in the online platform Cyberinfrastructure for Phylogenetic Research (CIPRES) (Miller et al. 2010). Tracer v1.5 was used to verify convergence (ESS > 200).

Divergence time

We estimated the divergence time between *C. ornatus* and *C. danae* mtDNA internal lineages using a Bayesian calibrated tree implemented in BEAST2. Other *Callinectes* species, *Arenaeus cribrarius* and *A. mexicanus* (outgroups) were included in our phylogenetic inferences (Table S1). The analyses were run using a birth–death speciation model and a relaxed log-normal molecular clock. Calibration point was set using the fossil of *Callinectes reticulatus* † Rathbun, 1918 (Late Oligocene), which is the oldest fossil in this genus (Luque et al. 2017). This information was included as a most recent common ancestor (mrca) prior at the node of the *Callinectes* monophyletic clade using an offset of 28 Mya and an exponential distribution. This point was chosen because there is no record of *Callinectes* from the Early Oligocene, so we used the date corresponding to the start of Late Oligocene. Due to this latter assumption, an exponential distribution is recommended because it creates a hard minimum bound (Ho and Phillips 2009). Two independent divergence time analyses were run in the online platform Cyberinfrastructure for Phylogenetic Research (CIPRES) with 100 million generations, sampling trees at every 10,000 generations, and assessed for convergence using Tracer v1.5. Ten per cent of the sampled trees were discarded as burn-in using TreeAnnotator (Drummond and Rambaut 2007), after combining runs using LogCombiner, and tree topology and divergence times were visualized in FigTree v1.4. (https://tree.bio.ed.ac.uk/software/figtree/).

Genetic distances

Genetic distances were calculated using Kimura-2-parameter (K2P) in MEGA 7 (Kumar et al. 2016). We only considered the COI dataset for this analysis because this is the traditional barcode gene (Hebert et al. 2003) and there are estimates of gap values available for crustaceans (Silva et al. 2011). We included other *Callinectes* sequences from GenBank to estimate intra- and interspecific genetic distances.

Approximate Bayesian computation (ABC) model testing

We identified two lineages within *C. ornatus*, suggesting that the Amazon-Orinoco plume is related to the intraspecific diversification of this species, but not to *C. danae* (see Results). Thus, we compared probable biogeographic scenarios that could explain this pattern for *C. ornatus* to better understand the role of the plume on this species. We performed an ABC analysis of historic demography to test the different evolutionary scenarios implemented in DIYABC (Cornuet et al. 2010) using concatenated COI and 16S genes. The scenarios tested were: (1) an ancestral population distributed along the western tropical Atlantic that split in two lineages after the inversion of the Amazon River caused by the Andean uplift; (2) The southern western Atlantic was the ancestral location of *C. ornatus*; from there it spread to the Caribbean; 3) the Caribbean was the ancestral location; from there it spread to the southern western Atlantic location (Fig. 2). These scenarios represent either a vicariant process (1) or a dispersion event (2 and 3). We considered that the dispersion events (2 and 3) were done by a small number of founders from the ancestral population, which experienced a population expansion after the event (Fig. 2). Summary statistics were chosen based on previous exploratory runs, and the best set was: variance of pairwise differences, mean of number of the rarest nucleotide at segregating sites, variance of numbers of the rarest nucleotide at segregating sites, and number of segregating sites between samples. Prior distributions of model parameters were set as: uniform distribution for both effective population size, ranging between 100 and 10,000; uniform distribution for divergence time between lineages (t2) ranging between 2,000,000 and 7,000,000, based on divergence time estimates using BEAST2 (see Results); uniform distribution for dispersion events (t1 in scenarios 2 and 3) ranging between 2,000,000 and 6,500,000; uniform distribution for founder population (for scenario 2 and 3) ranging between 10 and 10,000. We set conditions to t2 > t1, and founder populations < ancestral population. The

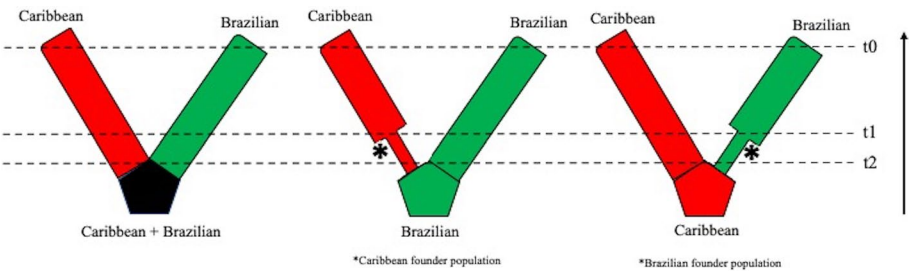


Fig. 2 Biogeographical scenarios drawn to test phylogeographic patterns in *Callinectes ornatus* (1) Vicariance scenario (2) and (3) Founder event scenario followed by population expansion in the receiver area. See the text for more information

mean mutation rate per site per generation was assigned using the prior uniform distribution of 10^{-8} to 10^{-7} adopted for mitochondrial markers (Cornuet et al. 2010), and an HKY mutation model. We simulated three million datasets (one million for each scenario). Model comparisons were based on polychotomous weighted logistic regression (Cornuet et al. 2010) on the 1% of simulated datasets that were closest to the observed data. The best scenario was chosen based on the highest significant posterior probability value, and on non-overlapping 95% confidence intervals. To have more confidence in our results, we evaluated the type I and II error rates. We simulated 500 pseudo-observed data per scenario by drawing parameter values from the prior distribution, and the posterior probabilities of each scenario were evaluated for every pseudo-observed dataset as described above. This step informed us of the probability that datasets simulated under the winner scenario were incorrectly assigned to another scenario (type I) and the probability that datasets simulated under another scenario were assigned to the winner scenario (type II).

Results

Sequence data

We generated 54 sequences of *Callinectes ornatus* and 60 sequences of *Callinectes danae* for both COI (570 bp/606 bp) and 16S (545 bp/436 bp). Sequences of both genes were obtained for each individual, and there were no missing data. The final concatenated sequence had 1115 bp (*Callinectes ornatus*) or 1042 bp (*C. danae*). The sequences generated in this manuscript have been submitted to the GenBank and the accession numbers are available in the Table S1.

Genetic structure analysis, genetic diversity and haplotype network

Callinectes ornatus full dataset had 39 haplotypes, 78 polymorphic sites, haplotype diversity of 0.957, and nucleotide diversity of 0.01360. *Callinectes danae* had 31 haplotypes, 34 polymorphic sites, a haplotype diversity of 0.822, and nucleotide diversity of 0.00145. The summary of each biogeographical province and panmictic population are indicated in Table 1.

The BAPS results suggested two clusters for *C. ornatus* (hereafter Caribbean and Brazilian ESUs) and one single cluster for *C. danae* (Fig. 3). Accordingly, the mtDNA haplotype network also indicates contrasting phylogeographic patterns between species (Fig. 4). The *C. ornatus* haplotype network was composed of two clusters separated by 32 mutation steps. These two clusters clearly separated Caribbean from Brazilian populations, and the only exception was a haplotype from the Caribbean that fell under the Brazilian group. In addition, the clusters had different shapes: The Caribbean cluster was reticulated whereas the Brazilian cluster was star-shaped. *Callinectes danae* haplotype network was composed of a single, star-shaped cluster.

Demographic history analysis

According to most neutrality tests, the Caribbean ESU of *C. ornatus* showed signals of population stability, whereas the Brazilian ESU showed signs of expansion. The

Table 1 Descriptive statistics of *Callinectes ornatus* and *C. danae* considering biogeographical provinces and panmictic populations. n: number of haplotypes; S: polymorphic sites; *h*: haplotypic diversity; π: nucleotide diversity

	Callinectes ornatus				*Callinectes danae*			
	n	S	*h*	π	n	S	*h*	π
Biogeographical provinces								
Tropical Northwestern Atlantic	7	48	0.773	0.00753	3	4	0.833	0.00224
North Brazil Shelf	2	1	1	0.0009	4	6	0.9	0.00231
Tropical Southwestern Atlantic	12	20	0.863	0.0026	18	19	0.897	0.00149
Warm Temperate Southwestern Atlantic	21	27	0.988	0.0031	10	11	0.727	0.00106
Panmictic population								
Caribbean	7	48	0.773	0.00753	31	34	0.822	0.00145
Brazilian	32	39	0.947	0.00265				

Fig. 3 BAPS-plot clustering result for *Callinectes ornatus* (left) and *C. danae* (right). Different colors represent different panmictic populations. Each vertical line in the BAPS-plot represent one specimen

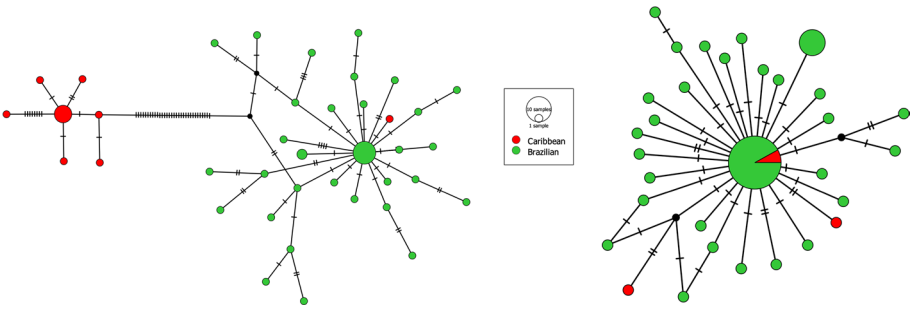


Fig. 4 Haplotype network result for *Callinectes ornatus* (left) and *C. danae* (right). The size of the network circles is proportional to the haplotype frequency. Different colors represent different panmictic populations

single ESU of *C. danae* showed signs of population expansion in all tests (Fig. 5, Table S3). Nevertheless, the BSP indicated population expansion in all cases but in different times. The *C. ornatus* Caribbean ESU seems to have experienced a long stability period, while the Brazilian ESU expanded approximately 50,000 years ago. *C. danae* ESU seems to have gone through expansion 25,000 years ago.

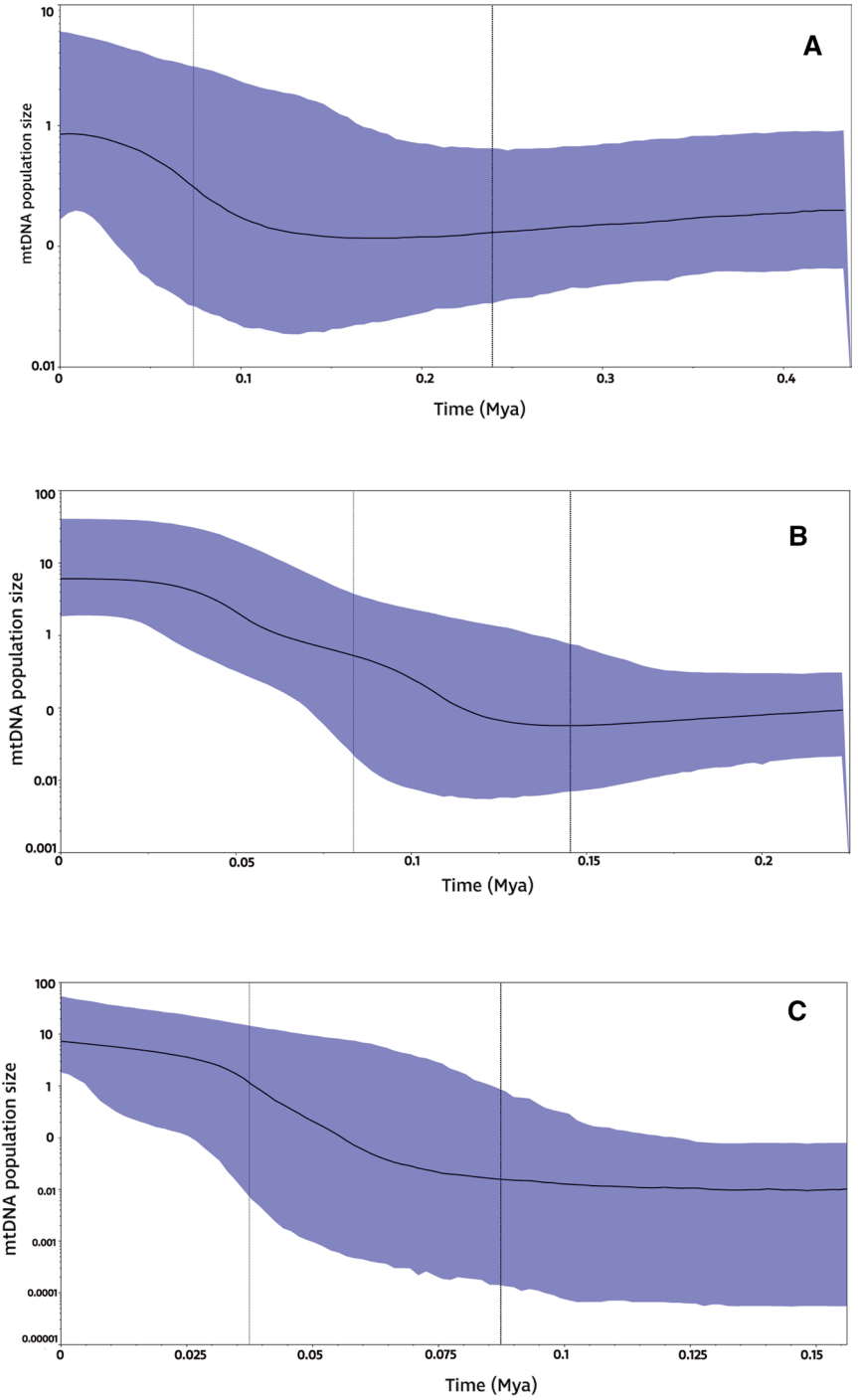


Fig. 5 Bayesian Skyline Plots (BSP) using concatenated COI/16S for **A** Caribbean *Callinectes ornatus* ESU, **B** Brazilian *Callinectes ornatus* ESU, **C** Single *Callinectes danae* ESU. The black line is the median estimate of the estimated effective population size. The blue area represents the upper and lower bounds of the 95% HPD interval

Divergence time

The Bayesian phylogenetic inference recovered the most common recent ancestor (MCRA) of *C. ornatus* clades 4.6 Mya (HPD 2.7–6.5 Mya), representing approximately when they diverged (Fig. 6). Even though we were not focusing on interspecific relationships, our analysis recovered *C. danae* x *C. similis* diverging recently at 1.3 Mya (HPD 0.6–1.9 Mya).

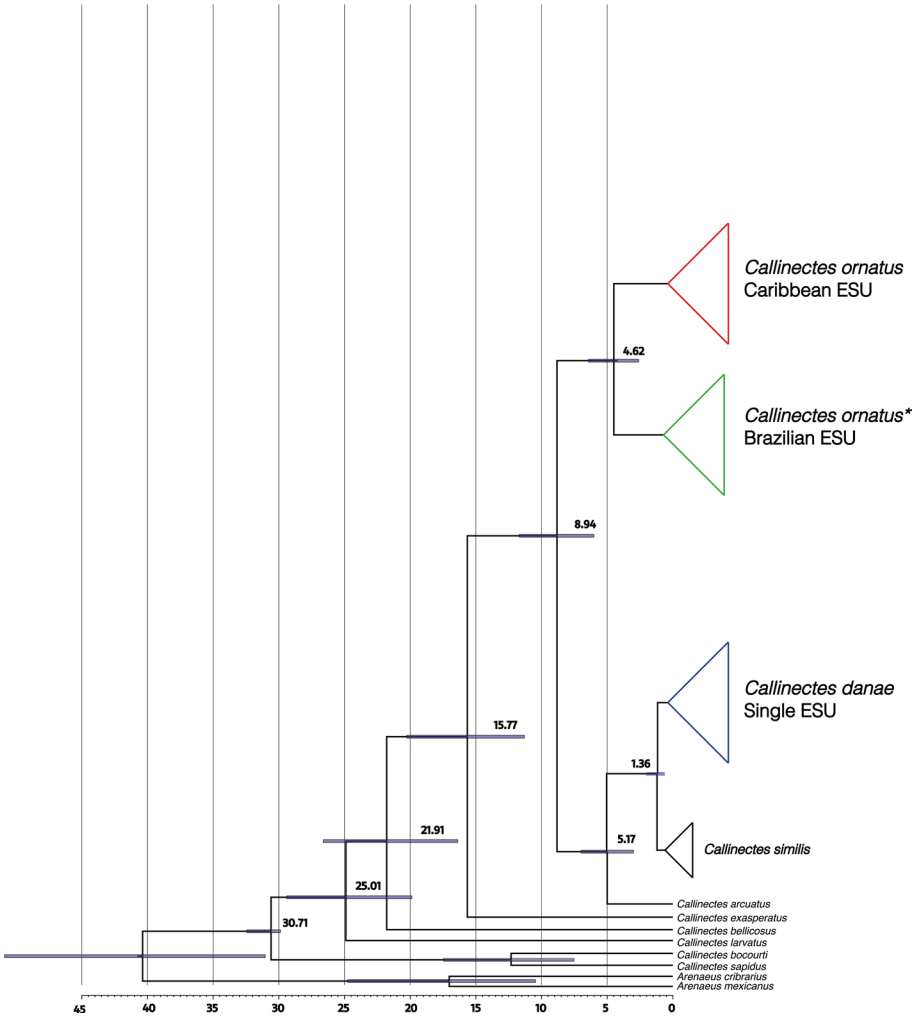


Fig. 6 Bayesian calibrated tree using concatenated COI/16S focusing on *Callinectes ornatus* and *C. danae*. Node bars represent the 95% HPD interval. Time divergence bar is in a Mya scale. *indicates that there is one individual from the Caribbean region within the Brazilian ESU. The image was edited to facilitate visualization of the results

Genetic distances

The pairwise genetic distance between *Callinectes* species ranged from 6.9 to 15.4% (Table S3). The distance between the Caribbean and Brazilian ESUs of *C. ornatus* was 4.7%, which is below the minimum distance between congeneric species we have found. However, the distance between *C. danae* and *C. similis* was 1.4%.

Approximate Bayesian computation (ABC) model testing

The ABC model testing identified scenario 3 as the most probable scenario (PP = 0.43; CI = 0.41–0.46) over scenario 1 (PP = 0.27; CI = 0.25–0.29) and scenario 2 (PP = 0.28; CI = 0.26–0.31) (Figure S1). This result supports our hypothesis testing of a dispersion event from Caribbean region to southeastern Atlantic after the Amazon-Orinoco plume establishment as the most probable scenario explaining *C. ornatus* phylogeographic pattern. We found moderately high rates of type I error, and low to moderate type II error rates (Table S4). Our results are probably limited by the use of mtDNA, which are prone to considerable stochastic variation in coalescent-based analysis (Cornuet et al. 2010). Even though these are not ideal rates, they still provide confidence in scenario 3.

Discussion

The use of biological traits during the formulation of hypotheses can offer new insights on patterns of genetic variation (Papadopoulou and Knowles 2016). Using co-distributed, closely related species as models we were able to isolate an important trait (salinity tolerance) and test its effects on genetic variation, under the scenario of a likely freshwater barrier. Here, *C. ornatus* and *C. danae* had contrasting phylogeographic structures and demographic histories, despite their close phylogenetic relationship. The genetic variation could not be explained by the long PLD. Thus, corroborating our hypothesis, the difference in salinity tolerance seems to be the key trait influencing the evolution of intraspecific lineages. Therefore, we propose that *C. ornatus* evolutionary history was driven by the Amazon-Orinoco plume, whereas *C. danae* evolutionary history was driven by the other factors we will discuss below. Our results contribute to the understanding of the biological traits that may be driving the diversification of western Atlantic species.

The role of the Amazon-Orinoco plume

The lack of well-defined patterns of genetic variation in populations found along the tropical Atlantic has been constantly discussed (Joyeux et al. 2001). Even though there are well-known barriers (Briggs and Bowen 2013), these can have an effective, null or variable role on phylogeographic patterns, depending on the species (Rocha and Bowen 2008). The Andean uplift affected markedly South American environments, and consequently, the western tropical Atlantic. The inversion of the Amazon river flow dated back to ~ 12 Mya (Hoorn et al. 1996, 2010), led to a great freshwater and sediment inflow into the Atlantic ocean, originating the Amazon-Orinoco plume. The plume had major effects on marine biota, especially after the Amazon River became

fully established around ~ 7 Mya (Hoorn et al. 1996, 2010). Using an explicit trait-based hypothesis approach, and comparing two congeneric species, we confirmed that the Amazon-Orinoco plume acted as a barrier for the species less tolerant to low salinity, i. e., *C. ornatus*.

Our mtDNA results indicated that *C. ornatus* is composed of two evolutionary significant units (ESUs sensu Moritz 1994): Caribbean and Brazilian. This division has also been found in other crabs (Laurenzano et al. 2016), polychaetes (Nunes et al. 2017), and fishes (Silva et al. 2018). Our divergence analysis estimated that these lineages diverged ~ 4.5 Mya, after the full formation of the Amazon River and consequent establishment of the Amazon-Orinoco plume as a barrier (Hoorn et al. 1996, 2010; Mandai et al. 2018). Our expectations were that a vicariant event was the determinant process leading to these two ESUs, as showed for other marine species (Boehm et al. 2013; Trovant et al. 2016). However, by means of ABC modelling we showed that this divergence resulted from a dispersion event of a few individuals from the Caribbean to the southeastern Atlantic, followed by a demographic expansion typical of founder events, and not by a vicariant process. We must take this result with caution since it did not present ideal support values, probably due to restrictions associated with the use of mtDNA in coalescent-based analysis (Cornuet et al. 2010). As a single locus, mtDNA is subjected to its own properties during the coalescent process. However, the Caribbean was already proposed as the ancestral area of *Callinectes* (Williams 1974), which supports our findings. Future analysis using high resolution markers might also confirm this scenario.

Commonly, the absence of gene flow would prevent the occurrence of shared haplotypes, and genetic drift could lead to complete divergence between regions, causing reciprocally monophyletic lineages and speciation (Avise 2000). This scenario would speculate that *C. ornatus* is in fact two taxonomic entities, and actually, many marine taxa have geminate species composed of Caribbean and Brazilian groups (Rocha 2003; Negri et al. 2014). However, our two *C. ornatus* groups are not morphologically different at least on basis of the commonly used taxonomic characters (Williams 1974; Santos 2007) and lack a molecular divergence gap large enough to indicate two species (Lefébure et al. 2006; Silva et al. 2011). Reinforcing this affirmation, there was a single specimen from French Antilles (Caribbean region) placed within the Brazilian group. This indicates that complete divergence has not taken place or that rare migrants may overpass the permeable Amazon-Orinoco plume.

Callinectes danae, on the other hand, and according to our predictions, is composed of a single ESU. This species has many physiological characteristics that ensure a successful dispersion across freshwater barriers. For instance, it is euryhaline and commonly found in estuaries and other low salinity areas (Shumway 1983; Mantelatto and Fransozo 2000) and not affected by freshwater outflow. Its long PLD (Costlow and Bookhout 1959; Bookhout and Costlow 1977) may promote the gene flow along the distribution range. In addition, *C. danae* has a special osmoregulatory ability (see Masui et al. 2002; Leone et al. 2005; Garçon et al. 2007 for review), which makes it more resistant to salinity variation than *C. ornatus*.

Our results indicate that, at least in marine taxa, the larval dispersal potential alone is not an enough predictor of phylogeographic structure (Weersing and Toonen 2009). However, when a species is tolerant to variation in abiotic conditions and it is not affected by biotic interactions, long pelagic dispersion can indeed guarantee the gene flow throughout the distributional range (Shanks 2009), as observed in *C. danae*. Many species occur in both Brazilian and Caribbean waters (Floeter et al. 2008), and do not form distinct genetic

groups (Laurenzano et al. 2013; Buranelli et al. 2019). Therefore, the role of the Amazon-Orinoco Plume on diversification is not idiosyncratic but depends on specific biological traits, such as salinity tolerance (Robertson et al. 2006; Hodge and Bellwood 2016).

Populations fluctuations through time

The demographic history of these species is also related with salinity tolerance. *Callinectes danae*, which occurs in estuarine and shallow coastal areas, showed signals of population expansion ~ 25,000 years ago, after the last glacial maxima (LGM) (Hewitt 2000). Other marine coastal invertebrates experienced expansion during this period, such as other Crustacea, Mollusca and Echinodermata species (Hellberg 2009; Marko et al. 2010). Paleoclimate changes had great influence on the sea level, affecting ocean currents, habitat extent, and biotic dynamics (Maggs et al. 2008; Marko et al. 2010; Briggs and Bowen 2013). *C. danae* may have undergone expansion because of the sea level rise after LGM, that increased coastal and estuarine areas, influencing its population size (Toms et al. 2014). According to our results the Brazilian *C. ornatus* group expanded before the LGM, and similar pre-LGM expansions have already been recorded in other coastal marine species (Marko et al. 2010). We may interpret the population expansion of *C. ornatus* as an indication that this species was less affected by LGM changes, compared to *C. danae.* Although *C. ornatus* occurs in shallow waters, they can migrate to deeper areas (Andrade et al. 2014). Thus, habitat changes during glaciations may not have affected *C. ornatus* populations so harshly. Contrastingly, the Caribbean population seems to have been stable through time or marked by a slightly recent expansion (BSP). The Caribbean is considered a center of biodiversity origin as well as accumulation (Rocha et al. 2008), and climate stability is one of the hypotheses raised to explain this feature (Fraser and Currie 1996; Mittelbach et al. 2007). Tropical stability would prevent populations from undergoing local extinctions due to paleoclimate changes and could maintain their population sizes stable through time (Carr et al. 2002). This result is reinforced by our estimation that the Caribbean is the ancestral area of *C. ornatus*.

Understanding how genetic variation is organized in space is also important considering current global changes and future scenarios. High-confidence predictions indicate increasing temperatures, acidification, and salinity on the Ocean (IPCC 2019). These abiotic parameters are already known to affect the invasion potential of the invasive crab *Carcinus maenas* (Compton et al. 2010), the egg size and embryonic development of *Neohelice granulata* (Giménez and Anger 2003), survivorship of *Callinectes sapidus* larvae (Giltz and Taylor 2017), and the whole physiology of *Callinectes danae* (Ramaglia et al. 2018). However, the consequences for genetic variation are not fully understood yet. It seems that species with wide tolerance ranges (*C. danae*) might be less affected by global changes, while those with narrow ranges (*C. ornatus*) might suffer from local population extinctions and changes in connectivity between populations. Future studies combining genetic data, species distribution modelling, traits data, could predict the impacts of global changes (e.g. Paz et al. 2019).

Taxonomic issues

Even though it was not under the scope of this study, we found an intriguing result related to the genetic divergence between *Callinectes danae* and *C. similis*. The pairwise COI divergence was 1.4%, which is an extremely low value for an interspecific gap (Lefébure et al. 2006; Silva et al. 2011). *Callinectes similis* occurs in sympatry with *C. danae* and

is restricted to the northwestern Atlantic. These species are very similar but recognized as two different species based on morphological (Williams 1974; Santos 2007), and molecular characters (Robles et al. 2007). If *C. danae* and *C. similis* are valid species, our two groups of *C. ornatus* may indicate that Caribbean and Brazilian ESU are each one a potential entity. However, the genetic distance between *C. ornatus* groups falls under the intraspecific gap known for decapods (Lefébure et al. 2006; Silva et al. 2011). Other decapod species, as well as other marine species (Derycke et al. 2010; Kieneke et al. 2012), for instance, also had higher divergence between groups separated geographically (Gouws et al. 2006). Additionally, Portunidae crabs seem to have higher divergence rates than other Brachyura families (Lefébure et al. 2006; Mantelatto et al. 2018; Spiridonov et al. 2014). Thus, we believe that the two *C. ornatus* units are isolated populations of the same species, and that occasional migration takes place. Regarding *C. danae* and *C. similis*, as far as we know, our study is the first one to include the DNA barcode region to infer their relationship, and due to the unusual finding, further studies are needed.

Conclusions

This study indicates that incorporating organismal biology information refines our hypothesis (Stewart et al. 2010; Paz et al. 2015; Papadopoulou and Knowles 2016) and helps the understanding of marine phylogeographic patterns. Considering the already identified marine barriers or biogeographical areas, we can target species that will accurately indicate which ecological traits are important drivers of genetic structure variation. Many studies focusing on terrestrial biota suggested (Prates et al. 2016) or identified ecological traits (e.g. habitat-use, Papadopoulou and Knowles 2015) that are important to explain phylogeographic patterns, but this kind of approach is still lacking for marine taxa. Even though studies focusing exclusively on mtDNA must be taken with caution, our results clearly indicate that ecological traits are more reliable predictors of phylogeographic structure than the standard dispersal potential approach. Further studies using high-resolution markers (e.g. SNPs) can lead to a deeper understanding of the Amazon-Orinoco Plume role in patterns of genetic variation. Nowadays, there are evidences that larval behavior, temperature and salinity tolerance, habitat fragmentation, anthropogenic activities and, notably, the diversity of life history strategies of marine organisms can affect current genetic distribution (Edmands and Potts 1997; Collin 2001; Nielsen and Kenchington 2001; Luttikhuizen et al. 2003; Baus et al. 2005; Kenchington et al. 2006; Crispo and Champman 2008). Therefore, there is an exciting range of possibilities to be explored in marine phylogeography.

Acknowledgements This research is resulted from the PhD project of Pedro A. Peres supported by São Paulo Research Foundation—FAPESP by a scholarship (2017/12376-6). All financial support was provided by grants from FAPESP (Temático Biota 2010/50188-8 and Biota INTERCRUSTA 2018/13685-5; Coleções Científicas 2009/54931-0; PROTAX 2016/50376-5), Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq (301359/2007-5; 473050/2007-2; 302748/2010-5; 471011/2011-8; PQ 302253/2019-0), and Coordenação de Aperfeiçoamento de Pessoal—CAPES—Código de Financiamento 001 (Ciências do Mar II—2005/2014–23038.004308/201414) granted to FLM. We are grateful to Gustav Paulay and John Slapcinsky (FLMNH), Cleverson Santos (MPEG) and all the colleagues of the Laboratory of Bioecology and Crustacean Systematics (LBSC) for their help with sampling, donation and lending specimens used herein. We are also grateful to Dr. Heather Bracken-Grissom for suggested improvements to the manuscript, to anonymous reviewers and editors for their suggestions and corrections that improved the manuscript during review process. The collections of species conducted in this study complied with current applicable state and federal laws of Brazil (permanent license to FLM for collection of Zoological Material No. 11777-1 MMA/IBAMA/SISBIO). The authors declare no conflicts of interest.

Authors contributions PAP: Conceptualization, Methodology, Formal analysis and Investigation, Writing – original draft, Writing – review and editing. FLM: Conceptualization, Writing – review and editing, Funding acquisition, Resources, Supervision.

References

Andrade LS, Bertini G, Fransozo V, Teixeira GM, de Paiva B-A, Fransozo A (2014) Differential occupation of habitat as a reproductive strategy of the blue crab *Callinectes ornatus* Ordway, 1968 (Crustacea: Decapoda). Mar Biodiver 44(1):27–36

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. J Biogeogr 28(7):819–825

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst 18(1):489–522

Ayre DJ, Hughes TP, Standish RJ (1997) Genetic differentiation, reproductive mode, and gene flow in the brooding coral *Pocillopora damicornis* along the Great Barrier Reef, Australia. Mar Ecol Prog Ser 159:175–187

Ayre DJ, Minchinton TE, Perrin C (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? Mol Ecol 18(9):1887–1903

Baus E, Darrock DJ, Bruford MW (2005) Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. Mol Ecol 14(11):3373–3382

Beheregaray LB (2008) Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol Ecol 17(17):3754–3774

Bell RC, Parra JL, Badjedjea G, Barej MF, Blackburn DC, Burger M, Channing A, Dehling JM, Greenbaum E, Gvoždík V, Kielgast J, Kusamba C, Lotters S, McLaughlin PJ, Nagy ZT, Rodel MO, Portik DM, Stuart BL, VanDerWal J, Zassi-Boulou AG, Zamudio KR (2017) Idiosyncratic responses to climate-driven forest fragmentation and marine incursions in reed frogs from Central Africa and the Gulf of Guinea Islands. Mol Ecol 26(19):5223–5244

Briggs JC, Bowen BW (2013) Marine shelf habitat: biogeography and evolution. J Biogeogr 40(6):1023–1035

Boehm JT, Woodall L, Teske PR, Lourie SA, Baldwin C, Waldman J, Hickerson M (2013) Marine dispersal and barriers drive Atlantic seahorse diversification. J Biogeogr 40(10):1839–1849

Bookhout CG, Costlow JD Jr (1977) Larval development of *Callinectes similis* reared in the laboratory. Bull Mar Sci 27(4):704–728

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10(4):e1003537

Buckley D (2009) Toward an organismal, integrative, and iterative phylogeography. BioEssays 31(7):784–793

Buranelli RC, Mantelatto FL (2017) Broad-ranging low genetic diversity among populations of the yellow finger marsh crab *Sesarma rectum* Randall, 1840 (Sesarmidae) revealed by DNA barcode. Crustaceana 90(7–10):845–864

Buranelli RC, Felder DL, Mantelatto FL (2019) Genetic diversity among populations of the Western Atlantic mangrove crab *Ucides cordatus* (Linnaeus, 1763)(Decapoda: Brachyura: Ocypodidae): evidence for panmixia and useful data for future management and conservation. J Crustacean Biol 39(4):386–395

Burton RS (1998) Intraspecific phylogeography across the Point Conception biogeographic boundary. Evolution 52(3):734–745

Carr MH, Anderson TW, Hixon MA (2002) Biodiversity, population regulation, and the stability of coral-reef fish communities. Proc Natl Acad Sci 99(17):11241–11245

Chacur MM, Negreiros-Fransozo ML (2001) Spatial and seasonal distributions of *Callinectes danae* (Decapoda, Portunidae) in Ubatuba Bay, São Paulo. Brazil J Crustacean Biol 21(2):414–425

Clement M, Posada DCKA, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9(10):1657–1659

Collin R (2001) The effects of mode of development on phylogeography and population structure of North Atlantic *Crepidula* (Gastropoda: Calyptraeidae). Mol Ecol 10(9):2249–2262

Compton TJ, Leathwick JR, Inglis GJ (2010) Thermogeography predicts the potential global range of the invasive European green crab (*Carcinus maenas*). Divers Distri 16(2):243–255
Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. Math Biosci 205(1):19–31
Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. Genetics 163(1):367–374
Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. Bioinformatics 20(15):2363–2369
Cornuet JM, Ravigné V, Estoup A (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC Bioinform 11(1):401
Costlow JD Jr, Bookhout CG (1959) The larval development of *Callinectes sapidus* Rathbun reared in the laboratory. Biol Bull 116(3):373–396
Crandall KA, Fitzpatrick JF Jr (1996) Crayfish molecular systematics: using a combination of procedures to estimate phylogeny. Syst Biol 45(1):1–26
Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15(7):290–295
Crispo E, Chapman LJ (2008) Population genetic structure across dissolved oxygen regimes in an African cichlid fish. Mol Ecol 17(9):2134–2148
Curtin TB (1986) Physical observations in the plume region of the Amazon River during peak discharge—II. Water masses. Cont Shelf Res 6(1–2):53–71
Dawson MN (2001) Phylogeography in coastal marine animals: a solution from California? J Biogeogr 28(6):723–736
Derycke S, Vanaverbeke J, Rigaux A, Backeljau T, Moens T (2010) Exploring the use of cytochrome oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. PLoS ONE 5(10):e13716
Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7(1):214
Drummond AJ, Rambaut A, Shapiro BETH, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22(5):1185–1192
Edmands S, Potts DC (1997) Population genetic structure in brooding sea anemones (*Epiactis* spp) with contrasting reproductive modes. Mar Biol 127(3):485–498
Floeter SR, Rocha LA, Robertson DR, Joyeux JC, Smith-Vaniz WF, Wirtz P, Edwards AJ, Barreiros JP, Ferreira CEL, Gasparini JL, Brito A, Falcon JM, Bowen BW, Bernardi G (2008) Atlantic reef fish biogeography and evolution. J Biogeogr 35(1):22–47
Fraser RH, Currie DJ (1996) The species richness-energy hypothesis in a system where historical factors are thought to prevail: coral reefs. Am Nat 148(1):138–159
Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol Ecol 10(12):2741–2752
Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147(2):915–925
Garçon DP, Masui DC, Mantelatto FL, McNamara JC, Furriel RPM, Leone FA (2007) K+ and NH4+ modulate gill (Na+, K+)-ATPase activity in the blue crab, *Callinectes ornatus*: fine tuning of ammonia excretion. Comp Biochem Physi A 147(1):145–155
Giltz SM, Taylor CM (2017) Reduced growth and survival in the larval blue crab *Callinectes sapidus* under predicted ocean acidification. J Shellfish Res 36(2):481–486
Giménez L, Anger K (2003) Larval performance in an estuarine crab, *Chasmagnathus granulata*, is a consequence of both larval and embryonic experience. Mar Ecol Prog Ser 249:251–264
Gouws G, Stewart BA, Daniels SR (2006) Phylogeographic structure of a freshwater crayfish (Decapoda: Parastacidae: *Cherax preissii*) in south-western Australia. Mar Freshwater Res 57(8):837–848
Hebert PD, Cywinska A, Ball SL, Dewaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond B Biol Sci 270(1512):313–321
Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Annu Rev Ecol Evol Syst 40:291–310
Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405(6789):907–913
Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after. Mol Phylogenet Evol 54(1):291–301
Ho SY, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58(3):367–380

Hoareau TB, Boissin E, Paulay G, Bruggemann JH (2013) The Southwestern Indian Ocean as a potential marine evolutionary hotspot: perspectives from comparative phylogeography of reef brittle-stars. J Biogeogr 40(11):2167–2179

Hodge JR, Bellwood DR (2016) The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. J Biogeogr 43(7):1324–1335

Hoorn C, Paxton CG, Crampton WG, Burgess P, Marshall LG, Lundberg JG, Räsänen ME, Linna AM (1996) Miocene deposits in the Amazonian foreland basin. Science 273:122–125

Hoorn C, Wesselingh FP, Ter Steege H, Bermudez MA, Mora A, Sevink J, Sanmartín I, Sanchez-Meseguer A, Anderson CL, Figueiredo JP, Jaramillo C, Riff D, Negri FR, Hooghiemstra H, Lundberg J, Stadler T, Särkinen T, Antonelli A (2010) Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927–931

IPCC, Shukla PR, Skea J, Calvo Buendia E, Masson-Delmotte V, Pörtner H-O, Roberts DC, Zhai P, Slade R, Connors S, van Diemen R, Ferrat M, Haughey E, Luz S, Neogi S, Pathak M, Petzold J, Portugal Pereira J, Vyas P, Huntley E, Kissick K, Belkacemi M, Malley J (2019) Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems

Joyeux JC, Floeter SR, Ferreira CEL, Gasparini JL (2001) Biogeography of tropical reef fishes: the South Atlantic puzzle. J Biogeogr 28(7):831–841

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772–780

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647–1649

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. PLoS ONE 5(1):e8594

Kenchington EL, Patwary MU, Zouros E, Bird CJ (2006) Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). Mol Ecol 15(7):1781–1796

Keunecke KA, D'Incao F, Verani JR, Vianna M (2012) Reproductive strategies of two sympatric swimming crabs *Callinectes danae* and *Callinectes ornatus* (Crustacea: Portunidae) in an estuarine system, south-eastern Brazil. J Mar Biol Assoc UK 92:343–347

Kieneke A, Martínez Arbizu PM, Fontaneto D (2012) Spatially structured populations with a low level of cryptic diversity in European marine Gastrotricha. Mol Ecol 21(5):1239–1254

Knowlton N, Weigt LA, Solorzano LA, Mills DK, Bermingham E (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260(5114):1629–1632

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874

Laurenzano C, Mantelatto FL, Schubart CD (2013) South American homogeneity versus Caribbean heterogeneity: population genetic structure of the western Atlantic fiddler crab *Uca rapax* (Brachyura, Ocypodidae). J Exp Mar Biol Ecol 449:22–27

Laurenzano C, Costa TM, Schubart CD (2016) Contrasting patterns of clinal genetic diversity and potential colonization pathways in two species of western Atlantic fiddler crabs. PLoS ONE 11(11):e0166518

Leigh JW, Bryant D (2015) popart: full-feature software for haplotype network construction. Methods Ecol Evol 6(9):1110–1116

Leone FA, Furriel RPM, McNamara JC, Mantelatto FL, Masui DC, Alves LR, Gonçalves RR, Garçon DP (2005) (Na+, K+)-ATPase from crustacean gill microsomes: a molecular marker to evaluate adaptation to biotopes of different salinity. Trends Comp Biochem Physiol 11:1–15

Lefébure T, Douady CJ, Gouy M, Gibert J (2006) Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Mol Phylogenet Evol 40(2):435–447

Luttikhuizen PC, Drent J, Baker AJ (2003) Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. Mol Ecol 12(8):2215–2229

Luque J, Schweitzer CE, Santana W, Portell RW, Vega FJ, Klompmaker AA (2017) Checklist of fossil decapod crustaceans from tropical America. Part I: Anomura Brachyura. Nauplius 25:e2017025

Magalhães T, Robles R, Felder DL, Mantelatto FL (2016) Integrative taxonomic study of the purse crab genus *Persephona* Leach, 1817 (Brachyura: Leucosiidae): combining morphology and molecular data. PLoS ONE 11(4):e0152627

Maggs CA, Castilho R, Foltz D, Henzler C, Jolly MT, Kelly J, Olsen J, Perez KE, Stam W, Väinölä R, Viard F (2008) Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. Ecology 89:S108–S122

Mandai SS, Buranelli RC, Schubart CD, Mantelatto FL (2018) Phylogenetic and phylogeographic inferences based on two DNA markers reveal geographic structure of the orangeclaw hermit crab *Calcinus tibicen* (Anomura: Diogenidae) in the western Atlantic. Mar Biol Res 14(6):565–580

Mantelatto FL (2000) Allocation of the portunid crab *Callinectes ornatus* (Decapoda: Brachyura) in the Ubatuba Bay, northern coast of São Paulo State, Brazil. Crustac Issues 12:431–443

Mantelatto FL, Fransozo A (2000) Brachyuran community in Ubatuba bay, northern coast of São Paulo state. Brazil J Shellfish Res 19(2):701–710

Mantelatto FL, Robles R, Biagi R, Felder DL (2006) Molecular analysis of the taxonomic and distributional status for the hermit crab genera *Loxopagurus* Forest, 1964 and *Isocheles* Stimpson, 1858 (Decapoda, Anomura, Diogenidae). Zoosystema 28(2):495–506

Mantelatto FL, Reigada AL, Gatti AC, Cuesta JA (2014) Morphology of the first zoeal stages of five species of the portunid genus *Callinectes* (Decapoda, Brachyura) hatched at the laboratory. An Acad Bras Ciênc 86(2):755–768

Mantelatto FL, Robles R, Wehrtmann IS, Schubart CD, Felder DL (2018) New insights into the molecular phylogeny of the swimming crabs of the genera *Portunus* Weber, 1795 and *Achelous* De Haan, 1833 (Brachyura: Portunidae) of the Americas. J Crustacean Biol 38(2):190–197

Marko PB, Hoffman JM, Emme SA, McGovern TM, Keever CC, Nicole Cox L (2010) The 'Expansion–Contraction'model of Pleistocene biogeography: rocky shores suffer a sea change? Mol Ecol 19(1):146–169

Masui DC, Furriel RPM, McNamara JC, Mantelatto FLM, Leone FA (2002) Modulation by ammonium ions of gill microsomal (Na+, K+)-ATPase in the swimming crab *Callinectes danae*: a possible mechanism for regulation of ammonia excretion. Comp Biochem Physi 132(4):471–482

Miller SA, Dykes DD, Polesky HFRN (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16(3):1215

Miller MA, Pfeiffer W, Schwartz T (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 gateway computing environments workshop (GCE) (pp 1–8).

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA, McCain CM (2007) Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol Lett 10(4):315–331

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. Trends Ecol Evol 9(10):373–375

Moura RL, Amado-Filho GM, Moraes FC, Brasileiro PS, Salomon PS, Mahiques MM, Bastos AC, Almeida MG, Silva JM, Araujo BF, Brito FP (2016) An extensive reef system at the Amazon River mouth. Sci Adv 2(4):e1501252

Negreiros-Fransozo ML, Fransozo A (1995) On the distribution of *Callinectes ornatus* Ordway, 1863 and *Callinectes danae* Smith, 1869 (Brachyura, Portunidae) in the Fortaleza Bay, Ubatuba, Brazil. Iheringia, Ser Zool 79:13–25

Negri M, Lemaitre R, Mantelatto FL (2014) Molecular and Morphological Resurrection of *Clibanarius symmetricus* a Cryptic Species Hiding Under the Name for the "Thinstripe" Hermit Crab *C. vittatus* (Decapoda: Anomura: Diogenidae). J Crustacean Biol 34(6):848–861

Negri M, Schubart CD, Mantelatto FL (2018) Tracing the introduction history of the invasive swimming crab *Charybdis hellerii* (A. Milne-Edwards 1867) in the Western Atlantic: evidences of high genetic diversity and multiple introductions. Biol Invasions 20(7):1771–1798

Nielsen EE, Kenchington E (2001) A new approach to prioritizing marine fish and shellfish populations for conservation. Fish Fish 2(4):328–343

Norse EA (1978) An experimental gradient analysis: hyposalinity as an "upstress" distributional determinant for Caribbean portunid crabs. Biol Bull 155(3):586–598

Nunes FL, Van Wormhoudt A, Faroni-Perez L, Fournier J (2017) Phylogeography of the reef-building polychaetes of the genus *Phragmatopoma* in the western Atlantic Region. J Biogeogr 44(7):1612–1625

Oliveira CMCA, Terossi M, Mantelatto FL (2019) Phylogeographic structuring of the amphidromous shrimp *Atya scabra* (Crustacea, Decapoda, Atyidae) unveiled by range-wide mitochondrial DNA sampling. Mar Freshwater Res 70(8):1078–1093

Pailler K, Bourlès B, Gouriou Y (1999) The barrier layer in the western tropical Atlantic Ocean. Geophy Res Lett 26(14):2069–2072

Palero F, Abelló P, Macpherson E, Gristina M, Pascual M (2008) Phylogeography of the European spiny lobster (*Palinurus elephas*): influence of current oceanographical features and historical processes. Mol Phylogenet Evol 48(2):708–717

Palumbi SR (1992) Marine speciation on a small planet. Trends Ecol Evol 7(4):114–118
Papadopoulou A, Knowles LL (2015) Species-specific responses to island connectivity cycles: refined models for testing phylogeographic concordance across a Mediterranean Pleistocene Aggregate Island Complex. Mol Ecol 24(16):4252–4268
Papadopoulou A, Knowles LL (2016) Toward a paradigm shift in comparative phylogeography driven by trait-based hypotheses. Proc Natl Acad Sci 113(29):8018–8024
Pârvulescu L, Pérez-Moreno JL, Panaiotu C, Drăguț L, Schrimpf A, Popovici ID, Zaharia C, Weiperth A, Gál B, Schubart CD, Bracken-Grissom H (2019) A journey on plate tectonics sheds light on European crayfish phylogeography. Ecol Evol 9(4):1957–1971
Paz A, Ibáñez R, Lips KR, Crawford AJ (2015) Testing the role of ecology and life history in structuring genetic variation across a landscape: a trait-based phylogeographic approach. Mol Ecol 24(14):3723–3737
Paz A, Spanos Z, Brown JL, Lyra M, Haddad C, Rodrigues M, Carnaval A (2019) Phylogeography of Atlantic Forest glassfrogs (Vitreorana): when geography, climate dynamics and rivers matter. Heredity 122(5):545
Peres PA, Lopes M, Negri M, Robles R, Santos CRM, Mantelatto FL (2020) Lack of population genetic structure among Brazilian populations of *Callinectes danae* (Brachyura: Portunidae): implication for management and conservation. Reg Stud Mar Sci. https://doi.org/10.1016/j.rsma.2020.101336
Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25(7):1253–1256
Prates I, Xue AT, Brown JL, Alvarado-Serrano DF, Rodrigues MT, Hickerson MJ, Carnaval AC (2016) Inferring responses to climate dynamics from historical demography in neotropical forest lizards. Proc Natl Acad Sci 113(29):7978–7985
Ramaglia AC, de Castro LM, Augusto A (2018) Effects of ocean acidification and salinity variations on the physiology of osmoregulating and osmoconforming crustaceans. J Comp Physiol B 188(5):729–738
Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Mol Biol Evol 19(12):2092–2100
Robertson DR, Karg F, de Moura RL, Victor BC, Bernardi G (2006) Mechanisms of speciation and faunal enrichment in Atlantic parrotfishes. Mol Phylogenet Evol 40(3):795–807
Robles R, Schubart CD, Conde JE, Carmona-Suárez C, Alvarez F, Villalobos JL, Felder DL (2007) Molecular phylogeny of the American *Callinectes* Stimpson, 1860 (Brachyura: Portunidae), based on two partial mitochondrial genes. Mar Biol 150(6):1265–1274
Rocha LA (2003) Patterns of distribution and processes of speciation in Brazilian reef fishes. J Biogeogr 30(8):1161–1171
Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. J Fish Biol 72(5):1101–1121
Rocha LA, Bass AL, Robertson DR, Bowen BW (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). Mol Ecol 11(2):243–251
Rocha LA, Rocha CR, Robertson DR, Bowen BW (2008) Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. BMC Evol Biol 8(1):157
Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nat Rev Genet 3(5):380
Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15(2):174–175
Santos CRM (2007) Biogeografia, sistemática e filogenia de Portunidae Rafinesque, 1815 do Oceano Atlântico Ocidental (Crustacea, Decapoda: ênfase em *Callinectes* Stimpson, 1860 e *Portunus* Weber, 1795. Dissertation. Fundação Universidade Federal do Rio Grande
Schubart CD, Huber MGJ (2006) Genetic comparisons of German populations of the stone crayfish, *Austropotamobius torrentium* (Crustacea: Astacidae). Bull Fr Pêche Piscic 380–381:1019–1028
Schubart CD, Neigel JE, Felder DL (2000) Use of the mitochondrial 16S rRNA gene for phylogenetic and population studies of Crustacea. Crustac Issues 12:817–830
Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. Biol Bull 216(3):373–385
Shumway SE (1983) Oxygen consumption and salinity tolerance in four Brazilian crabs. Crustaceana 44:76–82
Silva JM, Creer S, Dos Santos A, Costa AC, Cunha MR, Costa FO, Carvalho GR (2011) Systematic and evolutionary insights derived from mtDNA COI barcode diversity in the Decapoda (Crustacea: Malacostraca). PLoS one 6(5):e19449
Silva D, Martins K, Oliveira J, da Silva R, Sampaio I, Schneider H, Gomes G (2018) Genetic differentiation in populations of lane snapper (*Lutjanus synagris*–Lutjanidae) from Western Atlantic as revealed by multilocus analysis. Fish Res 198:138–149

Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc Natl Acad Sci 105(36):13486–13491
Sotelo G, Morán P, Posada D (2009) Molecular phylogeny and biogeographic history of the European *Maja* spider crabs (Decapoda, Majidae). Mol Phylogenet Evol 53(1):314–319
Spalding MD, Fox HE, Allen GR, Davidson N, Ferdaña ZA, Finlayson MA, Halpern BS, Jorge MA, Lombana AL, Lourie SA, Martin KD (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. Bioscience 57(7):573–583
Spiridonov VA, Neretina TV, Schepetov D (2014) Morphological characterization and molecular phylogeny of Portunoidea Rafinesque, 1815 (Crustacea Brachyura): Implications for understanding evolution of swimming capacity and revision of the family-level classification. Zool Anz 253(5):404–429
Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. Proc Natl Acad Sci 277(1682):661–671
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–595
Teske PR, Papadopoulos I, Mmonwa KL, Matumba TG, McQuaid CD, Barker NP, Beheregaray LB (2011) Climate-driven genetic divergence of limpets with different life histories across a southeast African marine biogeographic disjunction: different processes, same outcome. Mol Ecol 20(23):5025–5041
Toms JA, Compton JS, Smale M, von der Heyden S (2014) Variation in palaeo-shorelines explains contemporary population genetic patterns of rocky shore species. Biol Lett 10(6):20140330
Trovant B, Basso NG, Orensanz JM, Lessa EP, Dincao F, Ruzzante DE (2016) Scorched mussels (*Brachidontes* spp Bivalvia: Mytilidae) from the tropical and warm-temperate southwestern Atlantic: the role of the Amazon River in their speciation. Ecol Evol 6(6):1778–1798
Waples RS (1991) Pacific salmon Oncorhynchus spp., and the definition of "species" under the Endangered Species Act. Mar Fish Rev 53(3):11–22
Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. Mar Ecol Prog Ser 393:1–12
Williams AB (1974) The swimming crabs of the genus *Callinectes*. Fish Bull 72(3):685–798

Affiliations

Pedro A. Peres[1] · **Fernando L. Mantelatto**[1]

Fernando L. Mantelatto
flmantel@usp.br

[1] Laboratory of Bioecology and Crustacean Systematics (LBSC), Graduate Program in Comparative Biology, Department of Biology, Faculty of Philosophy, Sciences and Letters at Ribeirão Preto (FFCLRP), University of São Paulo (USP), Av. Bandeirantes 3900, Ribeirão Preto, São Paulo 14040-901, Brazil

Supplementary Material

Salinity tolerance explains the contrasting phylogeographic patterns of two swimming crabs species along the Tropical Western Atlantic.

Pedro A. Peres[1] and Fernando L. Mantelatto[1]

1Laboratory of Bioecology and Crustacean Systematics (LBSC), Graduate Program in Comparative Biology, Department of Biology, Faculty of Philosophy, Sciences and Letters at Ribeirão Preto (FFCLRP), University of São Paulo (USP), Av. Bandeirantes 3900, 14040-901, Ribeirão Preto, São Paulo, Brazil. FLM – flmantel@usp.br

Corresponding author: pedro.peres27@gmail.com

Table S1. Number of individuals of *Callinectes ornatus* and *Callinectes danae* analyzed per biogeographical province and GenBank accession number (GB). TNA: Tropical Northwestern Atlantic; NBS: North Brazil Shelf; TSA: Tropical Southwestern Atlantic; WTSA: Warm Temperate Southern Atlantic; TNP: Temperate Northern Pacific; TEP: Tropical Eastern Pacific. Bio. Reg: Biogeographical Region following Spalding et al 2007. Vouch. Coll.: Voucher Collection ID. *Latitude and Longitude are presented as indicated on the original Voucher Collection ID tag or GB, otherwise we indicated as not available (n/a).

Species	Biog. Reg.	Country	Locality	Latitude*	Longitude*	Vouch. Coll.	GB 16S	GB COI
Callinectes danae	TNA	USA	Florida Keys, Florida	24° 40' 22" N	81° 14' 26" W	FLMNH 11409	MT271768	MT272190
Callinectes danae	TNA	French Antilles	Saint Martin	18° 04' N	63° 01' W	FLMNH 32140	MT271769	MT272191
Callinectes danae	TNA	Colombia	Caribana Point	n/a	n/a	USNM:1071671	KY940142	KY940212
Callinectes danae	TNA	Colombia	Tinajones	n/a	n/a	USNM 1261632	KY940143	KY940213
Callinectes danae	NBS	Brazil	Calçoene, Amapá	n/a	n/a	CCDB 6112	KY940145	KY940215
Callinectes danae	NBS	Brazil	Salinópolis, Pará	n/a	n/a	LCD 2024	KY940123	KY940192

Callinectes danae	NBS	Brazil	Salinópolis, Pará	n/a	n/a	LCD 2024	KY940124	KY940193
Callinectes danae	NBS	Brazil	Belém, Pará	n/a	n/a	LCD s/n	KY940125	KY940194
Callinectes danae	NBS	Brazil	Belém, Pará	n/a	n/a	LCD s/n	KY940126	KY940195
Callinectes danae	TSA	Brazil	Beberibe, Ceará	04°10' 39.7" S	38° 05' 43" W	CCDB 2339	KY940088	KY940157
Callinectes danae	TSA	Brazil	Fortaleza, Ceará	n/a	n/a	LCD 02023	KY940121	KY940190
Callinectes danae	TSA	Brazil	Fortaleza, Ceará	n/a	n/a	LCD 02023	KY940122	KY940191
Callinectes danae	TSA	Brazil	Parnamirim, Rio Grande do Norte	05° 58' 58" S	35° 07' 21" W	CCDB 3387	KY940090	KY940159
Callinectes danae	TSA	Brazil	Parnamirim, Rio Grande do Norte	05° 58' 58" S	35° 07' 21" W	CCDB 3387	KY940091	KY940160
Callinectes danae	TSA	Brazil	Maceió, Alagoas	n/a	n/a	MZUSP 6626	KY940111	KY940180
Callinectes danae	TSA	Brazil	Ipojuca, Pernambuco	08º 33' 51"S	35º 01' 34." W	CCDB 4508	KY940104	KY940173
Callinectes danae	TSA	Brazil	Recife, Pernambuco	n/a	n/a	LCD 02965	KY940116	KY940185
Callinectes danae	TSA	Brazil	Recife, Pernambuco	n/a	n/a	LCD 02965	KY940117	KY940186
Callinectes danae	TSA	Brazil	Recife, Pernambuco	n/a	n/a	LCD 02965	KY940118	KY940187
Callinectes danae	TSA	Brazil	Recife, Pernambuco	n/a	n/a	LCD 02965	KY940119	KY940188
Callinectes danae	TSA	Brazil	Ipojuca, Pernambuco	08º 33' 51" S	35º 01' 34." W	CCDB 4508	KY940103	KY940172
Callinectes danae	TSA	Brazil	Entre Rios, Bahia	12° 24' 12" S	37° 53' 38" W	CCDB 289	KY940086	KY940155
Callinectes danae	TSA	Brazil	Porto Seguro, Bahia	16° 27' 32" S	39° 04' 08" W	CCDB 1446	KY940087	KY940156
Callinectes danae	TSA	Brazil	Salvador, Bahia	n/a	n/a	LCD 02019	KY940112	KY940181
Callinectes danae	TSA	Brazil	Salvador, Bahia	n/a	n/a	LCD 02019	KY940113	KY940182
Callinectes danae	TSA	Brazil	Salvador, Bahia	n/a	n/a	LCD 02019	KY940114	KY940183
Callinectes danae	TSA	Brazil	Salvador, Bahia	n/a	n/a	LCD 02019	KY940115	KY940184
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02012	KY940105	KY940174
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02013	KY940106	KY940175
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02015	KY940107	KY940176
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02012	KY940108	KY940177
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02012	KY940109	KY940178
Callinectes danae	WTSA	Brazil	Pontal do Paraná, Paraná	n/a	n/a	LCD 02931	KY940110	KY940179
Callinectes danae	WTSA	Brazil	Baía da Babitonga, Santa Catarina	n/a	n/a	LCD 3128	KY940127	KY940196
Callinectes danae	WTSA	Brazil	Baía da Babitonga, Santa Catarina	n/a	n/a	LCD 3009	KY940128	KY940197

Callinectes danae	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3930	KY940093	KY940162
Callinectes danae	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3930	KY940094	KY940163
Callinectes danae	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3930	KY940095	KY940164
Callinectes danae	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3930	KY940096	KY940165
Callinectes danae	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3930	KY940097	KY940166
Callinectes danae	WTSA	Brazil	Ilha dos Marinheiros, Rio Grande do Sul	n/a	n/a	LCD 1678	KY940134	KY940203
Callinectes danae	TSA	Brazil	Praia de Peruá, Espírito Santo	20º 50' 40" S	40º 43' 24" W	CCDB 4000	KY940098	KY940167
Callinectes danae	TSA	Brazil	Praia de Peruá, Espírito Santo	20º 50' 40" S	40º 43' 24" W	CCDB 4000	KY940099	KY940168
Callinectes danae	TSA	Brazil	Praia de Peruá, Espírito Santo	20º 50' 40" S	40º 43' 24" W	CCDB 4000	KY940100	KY940169
Callinectes danae	TSA	Brazil	Praia de Peruá, Espírito Santo	20º 50' 40" S	40º 43' 24" W	CCDB 4000	KY940101	KY940170
Callinectes danae	TSA	Brazil	Praia de Peruá, Espírito Santo	20º 50' 40" S	40º 43' 24" W	CCDB 4000	KY940102	KY940171
Callinectes danae	WTSA	Brazil	Ilha do Governador, Rio de Janeiro	n/a	n/a	LCD 02964	KY940129	KY940198
Callinectes danae	WTSA	Brazil	Ilha do Governador, Rio de Janeiro	n/a	n/a	LCD 02964	KY940130	KY940199
Callinectes danae	WTSA	Brazil	Ilha do Governador, Rio de Janeiro	n/a	n/a	LCD 02964	KY940131	KY940200
Callinectes danae	WTSA	Brazil	Ilha do Governador, Rio de Janeiro	n/a	n/a	LCD 02964	KY940132	KY940201
Callinectes danae	WTSA	Brazil	Ilha do Governador, Rio de Janeiro	n/a	n/a	LCD 02964	KY940133	KY940202
Callinectes danae	WTSA	Brazil	Ubatuba, São Paulo	23° 27' 24" S	45° 01' 20" W	CCDB 3445	KY940078	KY940147
Callinectes danae	WTSA	Brazil	Ubatuba, São Paulo	23° 27' 24" S	45° 01' 20" W	CCDB 3445	KY940079	KY940148
Callinectes danae	WTSA	Brazil	Ubatuba, São Paulo	23° 27' 24" S	45° 01' 20" W	CCDB 3445	KY940080	KY940149
Callinectes danae	WTSA	Brazil	Ubatuba, São Paulo	23° 27' 24" S	45° 01' 20" W	CCDB 3445	KY940146	KY940216
Callinectes danae	WTSA	Brazil	Cananéia, São Paulo	26° 10' 01" S	47° 54' 18" W	CCDB 3244	KY940081	KY940150
Callinectes danae	WTSA	Brazil	Cananéia, São Paulo	26° 10' 01" S	47° 54' 18" W	CCDB 3244	KY940082	KY940151

Callinectes danae	WTSA	Brazil	Cananéia, São Paulo	26° 10' 01" S	47° 54' 18" W	CCDB 3244	KY940083	KY940152
Callinectes danae	WTSA	Brazil	Cananéia, São Paulo	26° 10' 01" S	47° 54' 18" W	CCDB 3244	KY940084	KY940153
Callinectes danae	WTSA	Brazil	Cananéia, São Paulo	26° 10' 01" S	47° 54' 18" W	CCDB 3244	KY940085	KY940154
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271166	MT272154
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271168	MT272161
Callinectes ornatus	TSA	Brazil	Parnamirim, Rio Grande do Norte	5° 58' S	35° 07" W	CCDB 6105	KY940092	KY940161
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271167	MT272153
Callinectes ornatus	TSA	Brazil	Parnamirim, Rio Grande do Norte	5° 58' S	35° 07" W	CCDB 6105	MT271176	MT272139
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271189	MT272148
Callinectes ornatus	TSA	Brazil	Recife, Pernambuco	n/a	n/a	LCD 2965	KY940120	KY940189
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271157	MT272140
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271155	MT272162
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271184	MT272163
Callinectes ornatus	TSA	Brazil	Aracajú, Sergipe	11° 00' 70" S	37° 03' 06" W	CCDB 6130	MT271165	MT272141
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	6º 21' 11" S	35º 00' 1" W	CCDB 5421	MT271179	MT272132
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271188	MT272170
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271175	MT272155
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271186	MT272149
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271174	MT272142
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	6º 21' 11" S	35º 00' 1" W	CCDB 5421	MT271177	MT272146
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271158	MT272156
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271173	MT272157
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271172	MT272150
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271185	MT272143

Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271191	MT272144
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271183	MT272166
Callinectes ornatus	WTSA	Brazil	Camboriú, Santa Catarina	27º 00' 03" S	48º 37' 10" W	CCDB 4401	MT271169	MT272151
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	32º 10' 23" S	52º 06' 10" W	CCDB 3929	MT271171	MT272145
Callinectes ornatus	WTSA	Brazil	Rio Grande, Rio Grande do Sul	6º 21' 11" S	35º 00' 1" W	CCDB 5421	MT271178	MT272133
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 28' 31" S	44° 57' 18" W	CCDB 1537	MT271154	MT272164
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 358	MT271156	MT272135
Callinectes ornatus	WTSA	Brazil	Vitória, Espírito Santo	28° 18' 8" S	40° 17' 8" W	CCDB 4056	MT271182	MT272137
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 359	MT271161	MT272136
Callinectes ornatus	WTSA	Brazil	Vitória, Espírito Santo	28° 18' 8" S	40° 17' 8" W	CCDB 4056	MT271181	MT272138
Callinectes ornatus	WTSA	Brazil	Macaé, Rio de Janeiro	22° 25' 1" S	41° 44' 5" W	CCDB 4251	MT271192	MT272152
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23°26'10"S	45°01'36"W	CCDB 351	MT271190	MT272134
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 357	MT271187	MT272147
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	n/a	n/a	ULLZ 4178	AJ298186	MF490074
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 355	MT271163	MT272159
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' S	45° 09' W	CCDB 0126	MT271159	MT272165
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 352	MT271164	MT272168
Callinectes ornatus	WTSA	Brazil	Vitória, Espírito Santo	28° 18' 8" S	40° 17' 8" W	CCDB 4056	MT271180	MT272169
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' 10" S	45° 01' 36" W	CCDB 356	MT271162	MT272167
Callinectes ornatus	WTSA	Brazil	Ubatuba, São Paulo	23° 26' S	45° 09' W	CCDB 0126	MT271160	MT272160
Callinectes ornatus	TNA	French Antilles	Saint Martin	18° 6' 14" N	63° 14' 25" W	FLMNH 32103	MT271170	MT272158
Callinectes ornatus	TNA	Trinidad and Tobago	Trinidad Island	24° 40' 24" S	81° 14' 26" W	FLMNH 11249	MT271144	MT272175
Callinectes ornatus	TNA	Trinidad and Tobago	Trinidad Island	n/a	n/a	UF 11249	KY940135	KY94024
Callinectes ornatus	TNA	USA	Indian River Lagoon, Florida	29° 43' 02" N	81° 14' 49" W	FLMNH 34910	MT271148	MT272176
Callinectes ornatus	TNA	USA	Florida Keys, Florida	24° 40' 22" N	81° 14' 26" W	FLMNH 11409(4)	MT271146	MT272173
Callinectes ornatus	TNA	USA	Cape Sable, Florida	25° 2" N	81° 20" W	FLMNH 1476(2)	MT271149	MT272172
Callinectes ornatus	TNA	USA	Florida Keys, Florida	24° 40' 22" N	81° 14' 26" W	FLMNH 11409(2)	MT271147	MT272178

Callinectes ornatus	TNA	USA	Dry Tortugas, Florida	24° 26' 54" N	82° 17' 21" W	FLMNH 3982	MT271153	MT272177
Callinectes ornatus	TNA	USA	Florida Keys, Florida	24° 40' 22" N	81° 14' 26" W	FLMNH 11409(3)	MT271150	MT272180
Callinectes ornatus	TNA	USA	Biscayne Bay, Florida	25° 27' 06" N	80° 11' 49" W	FLMNH 26242	MT271151	MT272179
Callinectes ornatus	TNA	USA	Indian River Lagoon, Florida	29° 43' 02" N	81° 14' 49" W	FLMNH 34910	MT271152	MT272171
Callinectes ornatus	TNA	USA	Florida Keys, Florida	24° 31' 02" N	81° 58' 21" W	FLMNH 19804	MT271145	MT272174
Callinectes ornatus	TNA	USA	Cape Sable, Florida	25° 02" N	81° 20" W	FLMNH 1476(1)	MT271193	MT272181
Callinectes exasperatus	WTSA	Brazil	Cananeia, São Paulo	25°01'22" S	47°55'48" W	CCDB 802	KX06042	KX060222
Callinectes larvatus	TSA	Brazil	Parnamirim, Rio Grande do Norte	5° 58' 39" S	35° 07" 21" W	CCDB 6104	KY940089	KY940158
Callinectes bellicosus	TNP	Mexico	Gulf of California	n/a	n/a	ULLZ 4166	DQ407670	MG462555
Callinectes sapidus	WTSA	Brazil	São Vicente, São Paulo	n/a	n/a	CCDB 1680	JX123476	JX123452
Callinectes bocourtii	NBS	Brazil	Calcoene, Amapá	n/a	n/a	CCDB 6111	KY940144	KY940214
Callinectes similis	TNA	USA	Fort Pierce, Florida	n/a	n/a	UF 8023	KY940138	KY940207
Callinectes similis	TNA	USA	Fort Pierce, Florida	n/a	n/a	UF 8023	KY940205	KY940136
Callinectes similis	TNA	USA	Fort Pierce, Florida	n/a	n/a	UF 8023	KY940206	KY940137
Arenaeus cribrarius	WTSA	Brazil	Macaé, Rio de Janeiro	22° 25' 1" S	41° 44' 5" W	CCDB 3255	JX123461	JX123429
Arenaeus mexicanus	TEP	Costa Rica	Puntarenas	9° 44' 24" N	2° 50' 46" W	CCDB 2936	JX123471	JX123448

Table S2. Neutrality tests result for *Callinectes ornatus* and *C. danae*. Tests were performed using panmictic populations.

Callinectes ornatus				
	Brazilian		Caribbean	
	Mean	p-value	Mean	p-value
Tajima's D	-0.078	<0.001	-0.07	<0.01
Fu's Fs	-0.1	<0.001	0.148	0.09
R2	0.111	<0.001	0.168	0.41
Callinectes danae				
	Mean		p-value	
Tajima's D	-0.061		<0.001	
Fu's Fs	-0.059		<0.001	
R2	0.106		<0.001	

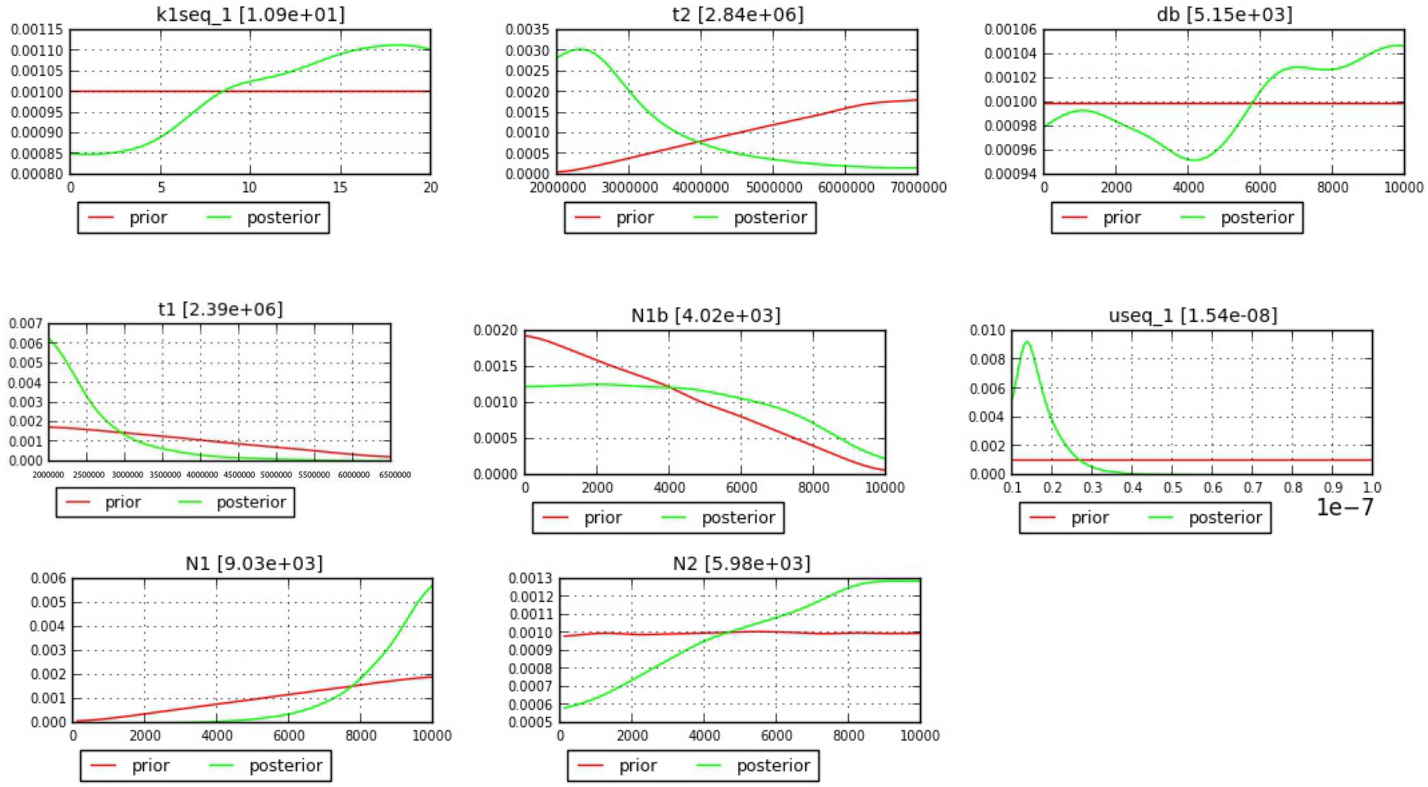
Table S3. Pairwise genetic distance (K2P) among *Callinectes* species using COI.

Species	1	2	3	4	5	6	7	8	9	10	11	12
1 *Callinectes danae*	0											
2 *Callinectes ornatus* BR	0.099	0										
3 *Callinectes ornatus* CR	0.108	0.049	0									
4 *Callinectes similis*	0.014	0.101	0.110	0								
5 *Callinectes arcuatus*	0.163	0.094	0.114	0.079	0							
6 *Callinectes exasperatus*	0.134	0.154	0.156	0.143	0.157	0						
7 *Callinectes larvatus*	0.163	0.156	0.169	0.156	0.169	0.180	0					
8 *Callinectes bellicosus*	0.174	0.164	0.154	0.178	0.164	0.175	0.169	0				
9 *Callinectes sapidus*	0.143	0.167	0.182	0.144	0.161	0.158	0.163	0.158	0			
10 *Callinectes bocourti*	0.160	0.153	0.157	0.159	0.174	0.170	0.145	0.156	0.118	0		
11 *Arenaeus cribrarius*	0.203	0.200	0.198	0.193	0.220	0.201	0.172	0.187	0.153	0.162	0	
12 *Arenaeus mexicanus*	0.209	0.198	0.205	0.205	0.223	0.209	0.179	0.187	0.196	0.174	0.147	0

Table S4. Results of model selection, type I and type II error rates for the scenarios estimated using DIYABC for *Callinectes ornatus* phylogeographic pattern.

Winner scenario	Type I error rate	Type II error rate		
		Scenario 1	Scenario 2	Scenario 3
Scenario 3 PP = 0.43 CI (0.41 – 0.46)	50.2%	30%	48.2%	-

Figure S1. DIYABC distribution plot for each estimated parameter.



Gmail

Pedro Peres <pedro.peres27@gmail.com>

Copyright Transfer Statement (CTS) for your article in Evolutionary Ecology (10057)

1 message

Springer <springerauthorquery@springeronline.com> Sun, Jun 7, 2020 at 9:38 AM
Reply-To: customerservice@springer.com
To: pedro.peres27@gmail.com

Copyright Transfer

07.06.2020

visit us at springer.com

Confirmation of your Copyright Transfer

Dear Author,

Please note: This e-mail is a confirmation of your copyright transfer and was sent to you only for your own records.

1. Publication
The copyright to this article, (including any supplementary information and graphic elements therein (e.g. illustrations, charts, moving images) (the 'Article'), is hereby assigned for good and valuable consideration to Springer Nature Switzerland AG (the 'Assignee'). Headings are for convenience only.

2. Grant of Rights
In consideration of the Assignee evaluating the Article for publication, the Author(s) grant the Assignee without limitation the exclusive (except as set out in clauses 3, 4 and 5 a) iv), assignable and sub-licensable right, unlimited in time and territory, to copy-edit, reproduce, publish, distribute, transmit, make available and store the Article, including abstracts thereof, in all forms of media of expression now known or developed in the future, including pre- and reprints, translations, photographic reproductions and extensions. Furthermore, to enable additional publishing services, such as promotion of the Article, the Author(s) grant the Assignee the right to use the Article (including the use of any graphic elements on a stand-alone basis) in whole or in part in electronic form, such as for display in databases or data networks (e.g. the Internet), or for print or download to stationary or portable devices. This includes interactive and multimedia use as well as posting the Article in full or in part or its abstract on social media, and the right to alter the Article to the extent necessary for such use. The Assignee may also let third parties share the Article in full or in part or its abstract on social media and may in this context sub-license the Article and its abstract to social media users. Author(s) grant to Assignee the right to re-license Article metadata without restriction (including but not limited to author name, title, abstract, citation, references, keywords and any additional information as determined by Assignee).

3. Self-Archiving
Author(s) are permitted to self-archive a pre-print and an author's accepted manuscript version of

their Article.

a. A pre-print is the author's version of the Article before peer-review has taken place ("Pre-Print"). Prior to acceptance for publication, Author(s) retain the right to make a Pre-Print of their Article available on any of the following: their own personal, self-maintained website; a legally compliant, non-commercial pre-print server such as but not limited to arXiv and bioRxiv. Once the Article has been published, the Author(s) should update the acknowledgement and provide a link to the definitive version on the publisher's website: "This is a pre-print of an article published in [insert journal title]. The final authenticated version is available online at: https://doi.org/[insert DOI]".
b. An Author's Accepted Manuscript (AAM) is the version accepted for publication in a journal following peer review but prior to copyediting and typesetting that can be made available under the following conditions:
    a. Author(s) retain the right to make an AAM of their Article available on their own personal, self-maintained website immediately on acceptance,
    b. Author(s) retain the right to make an AAM of their Article available for public release on any of the following 12 months after first publication ("Embargo Period"): their employer's internal website; their institutional and/or funder repositories. AAMs may also be deposited in such repositories immediately on acceptance, provided that they are not made publicly available until after the Embargo Period.

An acknowledgement in the following form should be included, together with a link to the published version on the publisher's website: "This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: http://dx.doi.org/[insert DOI]".

4. Authors' Retained Rights
Author(s) retain the following non-exclusive rights for the published version provided that, when reproducing the Article or extracts from it, the Author(s) acknowledge and reference first publication in the Journal:

a. to reuse graphic elements created by the Author(s) and contained in the Article, in presentations and other works created by them;
b. they and any academic institution where they work at the time may reproduce the Article for the purpose of course teaching (but not for inclusion in course pack material for onward sale by libraries and institutions); and
c. to reproduce, or to allow a third party Assignee to reproduce the Article in whole or in part in any printed volume (book or thesis) written by the Author(s).

5. Warranties
The Author(s) warrant and represent that:

a. (i) the Author(s) are the sole copyright owners or have been authorised by any additional copyright owner(s) to assign the rights defined in clause 2, (ii) the Article does not infringe any intellectual property rights (including without limitation copyright, database rights or trade mark rights) or other third party rights and no licence from or payments to a third party are required to publish the Article, (iii) the Article has not been previously published or licensed, (iv) if the Article contains material from other sources (e.g. illustrations, tables, text quotations), Author(s) have obtained written permissions to the extent necessary from the copyright holder(s), to license to the Assignee the same rights as set out in Clause 2 but on a non-exclusive basis and without the right to use any graphic elements on a stand-alone basis and have cited any such material correctly;
b. all of the facts contained in the Article are according to the current body of science true and accurate;
c. nothing in the Article is obscene, defamatory, violates any right of privacy or publicity, infringes any other human, personal or other rights of any person or entity or is otherwise unlawful and that informed consent to publish has been obtained for all research participants;
d. nothing in the Article infringes any duty of confidentiality which any of the Author(s) might owe to anyone else or violates any contract, express or implied, of any of the Author(s). All of the institutions in which work recorded in the Article was created or carried out have authorised and approved such research and publication; and
e. the signatory (the Author or the employer) who has signed this agreement has full right, power and authority to enter into this agreement on behalf of all of the Author(s).

6. Cooperation
The Author(s) shall cooperate fully with the Assignee in relation to any legal action that might arise from the publication of the Article, and the Author(s) shall give the Assignee access at reasonable times to any relevant accounts, documents and records within the power or control of the Author(s). The Author(s) agree that the distributing entity is intended to have the benefit of and shall have the right to enforce the terms of this agreement.

7. Author List
After signing, changes of authorship or the order of the authors listed will not be accepted unless formally approved in writing by the Assignee.

8. Edits & Corrections
The Author(s) agree(s) that the Assignee may retract the Article or publish a correction or other notice in relation to the Article if the Assignee considers in its reasonable opinion that such actions are appropriate from a legal, editorial or research integrity perspective.

This is an automated e-mail; please do not reply to this account. If you have any questions, please go to our help pages.

Thank you very much.

Kind regards,

Springer Author Services

Article Details

Journal title	Article title
Evolutionary Ecology	Salinity tolerance explains the contrasting phylogeographic patterns of two swimming crabs species along the Tropical Western Atlantic.
DOI	**Corresponding Author**
10.1007/s10682-020-10057-x	Pedro Peres
Copyright transferred to	**Transferred on**
Springer Nature Switzerland AG	Sun Jun 07 14:38:48 CEST 2020

Service Contacts

Springer Nature Customer Service Center

Tiergartenstr. 15-17
69121 Heidelberg
Germany

Springer New York, LCC

233 Spring Street
New York, NY 10013
USA
phone: +1 212 460 1500 or 800-SPRINGER

phone: +49 6221 345 0
fax: +49 6221 345 4229
customerservice@springernature.com

(Weekdays 8:30am - 5:30pm ET)
fax: +1 212-460-1700
customerservice@springernature.com

CHAPTER 2 (Manuscript *in prep.*)

Genomic analyses suggest incipient speciation in a widespread Tropical Atlantic swimming crab

Pedro A. Peres; Laura Timm; Heather Bracken-Grissom; Fernando L. Mantelatto

Introduction

The drivers and consequences of the temporal and spatial components of genetic diversity in the marine environment remain a challenge for ecologists and evolutionary biologists (Bowen et al., 2014). Although coastal and open ocean ecosystems harbor a considerable species diversity, little is known about these species compared to the terrestrial environment, particularly regarding tropical marine invertebrates (Beheregaray, 2008), including the speciose group of decapod crustaceans. Traditionally, marine systems have been considered more "open" due to the transport of particles by currents and the wide distribution of species, resulting in higher import and export rates among regions (Palumbi, 1992). Many marine species show indirect development with larval stages during their life cycle. The larvae are released by adults in the water and have the potential to be transported by ocean currents and reach distant locations (Hedgecock, 1986). In general, there is a positive association between the pelagic larval duration (PLD) and the distance traveled by the larvae (Shanks, 2009). Consequently, longer PLD would result in a lack of population differentiation, but this is not always true.

The association between PLD and population differentiation has been under debate in the last two decades. Although this is a prevalent association in many species (e.g., Reece et al., 2011), it has been shown to represent a weak or null correlation across marine animals (Weersing & Toonen, 2009). Indeed, many species show some kind of genetic structure across their distribution (Pelc et al., 2009; Kelly & Palumbi, 2010). Larval retention caused by local currents, larval behavior, and larval development rate variation depending on the region or adult's ecology is among the factors explaining why a long PLD not always result in connectivity among distant populations (White et al., 2010; Butler et al., 2011 Hedgecock & Pudovkin, 2011; Álvarez-Noriega et al., 2020; Timm et al., 2020). Thus, even

phylogenetically close and co-distributed species with similar PLD can also show contrasting genetic patterns (Eble et al., 2009).

Since DNA started to be used to investigate marine species, we find evidence for underestimation of the diversity (Knowlton, 1993), especially in cases where widespread species show high genetic structure or reveal to be, in fact, a complex of species (Gaither et al., 2010; Leasi et al., 2016; Álvarez-Campos et al., 2017). This happens due to the lack of or subtle morphological variation (Schubart et al., 2000) and cryptic barriers (Briggs & Bowen, 2013). In the marine environment, barriers are represented by the distance among regions or by abiotic breaks/gradients rather than by conspicuous barriers (e.g., mountains in the terrestrial environment) and can be easily overlooked (Briggs & Bowen, 2013)

With the advances of powerful molecular markers generated through next-generation sequencing (NGS), like SNPs, previously undetected population structure is being revealed. Studies investigating species showing long PLD, like the lobster *Homarus americanus* (PLD = 21 to 56 days) and the sea cucumber *Parastichopus californicus* (PLD = 30 to 120 days), employing RAD-seq and ddRAD-seq, respectively, found fine-scale structure despite the dispersal potential (Benestan et al., 2015; Xuereb et al., 2018). NGS approaches also opened a new venue to explore adaptive structure and clines in the marine environment. Using reduced representation libraries (RRL), it is possible to access variation across the genome and potentially capture protein-coding genes (Catchen et al., 2017). For instance, the sandgoby *Psammogobius knysnaensis* shows neutral genetic homogeneity along the South African coast, but it is composed of distinct groups when looking at markers related to thermal tolerance (Teske et al., 2019). In the Red Sea fish species *Amphiprion bicinctus*, an environmental transition from an oligotrophic area to a eutrophic area represents a genetic break (Saenz-Agudelo et al., 2015). The discovery of subtle or strong genetic differentiation boosted by the use of powerful markers raises the possibility for a deeper understanding of the diversification process in marine species at both intra- and interspecific levels. Neutral or adaptive structures maintained for long periods have the potential of generating new species (Seehausen et al., 2014; Kulmuni et al., 2020).

In the tropical Western Atlantic, we find the world's largest freshwater and sediment discharge into the ocean: The Amazon-Orinoco plume (Curtin, 1986). The plume is formed by the combination of the Amazon and Orinoco river mouths in the North part of South America. The beginning of the outflow and sediment deposition date back to the Miocene (10 MYA) due to the uplift of the Andes mountains, becoming established around 7 MYA (Hoorn et al., 1996; Hoorn et al., 2010). The Amazon-Orinoco plume represents an environmental barrier reaching 400 km off the coast and 30 m in depth (Pailler et al., 1999). However, the plume is considered a soft barrier because its effect depends on the species (Briggs & Bowen, 2013). We find evidence for genetic structure or sister species between both sides of the plume, so as no influence at all can be detected in some manatees, fish, annelids, mussels, and crustaceans (Joyeux et al., 2001; Tourinho et al., 2012; Trovant et al., 2016; Nunes et al., 2017; Silva et al., 2018; Luna et al., 2021). Species or groups of species distributed along the tropical Western Atlantic coast, including the Amazon-Orinoco plume, represent promising candidates for investigations on neutral and adaptive structure using NGS approaches. Many questions remain unanswered as we do not know if species showing genetic homogeneity might show adaptive structure, we do not have robust divergence time estimates resulted from hundreds of loci, or comparisons between markers generated from sanger sequencing vs. NGS, and others. However, we still lack studies under this framework or analyses species occurring on both sides of the plume employing NGS (but see Titus et al., 2019; Pedraza-Marrón et al., 2019)

The swimming crab *Callinectes ornatus* represents a potential model to investigate the diversification process in the tropical Western Atlantic and the effects of a soft barrier on the genetic structure of marine species. The species is widespread along the Western Atlantic, occurring in coastal waters (up to 75m) from south Brazil to North Carolina (USA) (Norse 1978; Mantelatto & Fransozo 2000). Despite high dispersal potential, the species show a strong genetic structure, composed of two separated groups: one north and one south the Amazon-Orinoco plume (Peres & Mantelatto, 2020). The authors argued that salinity tolerance plays a significant role in defining these two groups. Still, they were limited in their conclusion due to the use of a single locus (mtDNA) and could not confirm if

ongoing migration was occurring or if both species were completely isolated representing separate species. Therefore, the use of NGS approaches can elucidate this species lineage diversification and shed light on the processes acting upon Western Atlantic species.

Here, we combined available and novel mtDNA data (Peres & Mantelatto, 2020) with ddRAD-seq (Peterson et al., 2012) to explore the diversification patterns and processes within the swimming crab *Callinectes ornatus.* We investigated 1) the existence of fine-scale structure across Brazilian populations; 2) the differentiation level between populations from both sides of the Amazon-Orinoco plume; 3) the occurrence of loci under selection and adaptive structure; 4) the potential existence of recently diverged species. Our dataset also allowed us to investigate the differences and similarities between the types of molecular markers employed and the results generated using SNPs datasets built under different filtering settings. To the best of our knowledge, the work we present here represents the first investigation employing ddRAD-seq in a marine invertebrate species with distribution encompassing tropical north and south Western Atlantic.

Materials and Methods

Sampling and DNA extraction for mtDNA and ddRAD-seq

For the mtDNA analyses, we used the same cytochrome c oxidase subunit I (COI) dataset as in Peres & Mantelatto (2020). We expanded it with new sequences available on Genbank, plus new sequences generated and submitted to GenBank. For the ddRAD-seq analyses, we obtained 63 individuals of *C. ornatus* from the following collections: Crustacean Collection of the Department of Biology (CCDB), Faculty of Philosophy, Sciences and Letters at Ribeirão Preto (FFCLRP) of the University of São Paulo (USP), the Invertebrate Zoology Collection - Florida Museum of Natural History (FLMNH) of the University of Florida, and the Florida International University Crustacean Collection (FICC). Many of them were previously used in mtDNA analyses. Our sampling covers all

of the species range distribution, and for ddRAD-seq analyses, we considered the following populations: Caribbean (CR), Northeast Brazil (NE), Southeast Brazil (SE), and South Brazil (S) (Figure 1).

Genomic DNA was extracted from muscle tissues using the salt extraction method (Miller et al., 1988), with the modifications indicated in Mantelatto et al. (2006, 2018), or with the DNeasy Blood and Tissue Kit (Qiagen), following the protocol provided by the manufacturer.

Figure 1. Map of Western Atlantic showing *Callinectes ornatus* sampled populations. Circles represent the populations used for ddRAD-seq analyses. Red: Caribbean; Green: Northeastern Brazil; Blue: Southeastern Brazil; Yellow: South Brazil. Triangles represent COI new sequences added to Peres & Mantelatto (2020) dataset. Blue lines represent Amazon River and Orinoco River mouth, resulting the Amazon-Orinoco Plume. Arrows represent oceanic currents.

COI amplification and analyses

COI sequences were amplified using the primer COL6b/COH6 (Schubart & Huber 2006), and PCR cycles, DNA purification, and sequencing following Peres & Mantelatto (2020). Quality filtering, trimming, primer removal, and *denovo* assembling steps were performed in Geneious Prime 2020.2.4 (https://www.geneious.com). We checked for pseudogenes by translating the consensus sequences and checking for indels and stop codons, removing them when present (Song et al., 2008). We aligned the sequences using MAFFT v.7 (Katoh & Standley, 2013) and prepared a haplotype file in DnaSP v.6 (Rozas et al., 2017), which was used to access the relationship among haplotypes using a statistical parsimony network through the TCS method (Clement et al., 2000) implemented in PopArt 1.7(Leigh & Bryant, 2015).

ddRAD-seq library preparation and data processing

Double digest RADseq libraries were prepared according to the ddRADseq method (Peterson et al., 2012). DNA from all individuals was digested with a combination of NlaIII and NotI (New England Biolabs) after enzyme trials to determine the best enzyme set. Following digestion, custom barcoded adapters were ligated to the fragments and pooled into eight sublibraries. Each sublibrary was size selected (250 – 300 base pairs - bp) on a PippinPrep (SageScience). Size selected fragments were then amplified via PCR with Phusion Hi-Fidelity Polymerase (Thermo Scientific), which also incorporated indices (i7) and Illumina adapters into the fragments, allowing for pooling of sublibraries into the final library. Sequencing was done at the Genewiz Facility in South Plainfield, New Jersey, by an Illumina HiSeq 4000 (PE150).

Raw sequence files were quality-filtered, aligned, and assembled with the STACKS v2.3d (Rochette et al., 2019) on the Florida International University High-Performance Computing Cluster (HPCC). Reads were demultiplexed, cleaned, and quality-filtered with the process_radtags program.

We used the denovo_map.pl wrapper program to run ustacks, cstacks, sstacks, tsv2Bam, and populations modules on STACKS v2.3d. Exploratory analyses were run using a subset of our data to determine the best parameter set, following Rochette et al. (2017) suggestions to maximize SNPs calling and minimize error rates. Identical reads were aligned within each individual in ustacks, and putative paralogs were excluded by setting the maximum distance between reads in a stack to 5 (-M 5) and the minimum depth of coverage required to create a stack was set to 10 (-m 10). This last parameter was set because depth >10X effectively reduces genotype error (Fountain et al., 2016). Consensus reads were cataloged in cstacks, and the number of mismatches allowed between sample loci was set to 5 (-n 5). All putative loci were matched against the catalog with sstacks, and forward and reverse (paired-end) reads were assembled by tsv2Bam. Because not all of our individuals had high molecular weight DNA yields, which can impact downstream analyses (Peterson et al., 2012; Cumer et al., 2021), we decided to run a population module test to explore our data. We did not use any filtering on this exploratory analysis. Then, we used the function --missing-indv on VCFtools v. 0.1.17 (Danecek et al., 2011) to access the maximum number of SNPs retrieved for each individual. Individuals showing <200 SNPs were excluded (n = 18). We decided to design nine different datasets (Table 1) using different filtering approaches by changing population module settings (see below) to account for the effects of missing data and its impact on our conclusions.

The population module was used to generate a file of unlinked single nucleotide polymorphisms (SNPs) and for each dataset. Our strategy was to progressively exclude low coverage individuals and change the minimum percentage of individuals in a population required to process a locus for that population (-r) (Table 1). The datasets are composed of 30, 40, and 45 individuals respectively. Depending on the dataset, a SNP had to be present in 25% (-r 0.25), 50% (-r 0.5), or 75% (-r 0.75) of the individuals of a population to be called for that population. We set a minor allele frequency of 5% (--min-maf 0.05) for all datasets because low-frequency alleles can affect population structure inferences (Linck & Battey, 2018). Also, we set a maximum observed heterozygosity of 50% (--max-obs-het 0.5) because biallelic SNPs are expected to show heterozygosity <50%, preventing

multilocus contigs or paralogous loci to be included in our analyses (Hohenlohe et al., 2011; Willis et al., 2017; Gargiulo et al., 2020). Finally, one SNP was called per locus (-- write_random_snp) to generate a final alignment of unlinked SNP.

Table 1. Datasets used in population structure analyses. Each column show the number of individuals (n), the criterion used to exclude individuals, the number of individuals per population (CR: Caribbean, NE: Northeastern Brazil, SE: Southeastern Brazil, S: South Brazil), and the number of SNPs retained after applying different -r filters.

	n = 30 (<1000 SNPs excluded) CR (4), NE (7), SE (10), S (9)	n = 40 (<400 SNPs excluded) CR (6), NE (12), SE (11), S (11)	n = 45 (<200 SNPs excluded) CR (6), NE (12), SE (14), S (13)
-r 0.25	8202	7293	5559
-r 0.50	3937	2232	1887
-r 0.75	1519	682	324

Outlier detection

Prior to further analyses, we identified outlier SNPs (i.e., non-neutral SNPs, possibly under selection) using BayeScan v.2.1 (Foll & Gaggiotti, 2008) and PCAdapt (Luu et al., 2017). We used two different approaches because it is recommended to use multiple methods to identify non-neutral SNPs to reduce type 1 error (false positive) (Narum & Hess, 2011; Villemereuil et al., 2014; François et al., 2016). We ran both analyses for all datasets. BayeScan is a method to identify putative adaptive SNPs based on different allele frequencies among populations, and we performed it setting up prior odds to 10, iterations to 5000, and burn-in to 50000. Outlier loci were identified at a q-value (i.e., false discovery rate) of 0.05 and removed from the neutral-loci dataset. BayeScan was ran on the Florida International University High-Performance Computing Cluster (HPCC). PCAdapt implements a hierarchical method (not assuming populations a priori) based on principal component analysis that identifies SNPs excessively related to population structure, probably due to selection. We ran PCAdapt exploring twenty PCs (K = 20) to select the optimal K following Cattel's rule to retain the best K PC value, depending on

the dataset (Luu et al., 2017). These PC detect the SNPs most associated with population structure. We filtered putative non-neutral SNPs based on a q-value of 0.01. A SNPs was considered as non-neutral when both analyses across all datasets identified it. Hereafter, analyses considered a neutral- and a non-neutral dataset, and all SNPs under selection were removed from all neutral datasets. All outlier loci were subjected to a Basic Local Alignment Search Tool (BLAST) search to check if they match annotated sequences available in NCBI public database (Johnson et al., 2008). We optimize the search using the blastn algorithm and constraining the search set to organism:decapoda (taxid:6683) because the entry "brachyura" resulted in no hits.

Population structure

We calculated corrected pairwise-F_{ST} on GENODIVE v3.0 (Meirmans, 2020) with 999 permutations to access significance for all nine neutral datasets. We employed the Bayesian program STRUCTURE v2.3.4 (Pritchard et al., 2000) to test for population structure within the data. Seven K-values were tested (K = 8) 10 times each under the admixture model. Following a burn-in of 10,000 generations, 100,000 Markov Chain Monte Carlo generations ran. In STRUCTURE HARVESTER v0.6.94 (Earl, 2012), STRUCTURE results were collated, and ad hoc posterior probability models (Pritchard et al., 2000) and the Evanno method (Evanno et al., 2005) were used to infer the optimal K value. Both analyses were also done using the non-neutral dataset to access adaptive structure.

Divergence time, migration, and effective population size

We implemented the Generalized Phylogenetic Coalescent Sampler (G-PhoCS) to model the demographic history of *C. ornatus* (Gronau et al., 2011). G-PhoCS is based on the MCMCcoal model, which employs a multispecies coalescent framework to estimate divergence times and effective population sizes from multilocus sequence data (Yang 2002; Rannala & Yang, 2003), but additionally

allows for modeling gene flow between populations along with user-defined migration bands (Gronau et al., 2011). G-PhoCS uses a full-likelihood, coalescent model where the input data wasthe entire sequence for each ddRAD locus. We decided to use a subset of our individuals and loci due to computational time to convergence. We kept seven individuals from the CR group and nine individuals from the BR group (3 from each population) because it is a sufficient number to achieve reliable results (I. Gronau personal communication). We ran the population module using this smaller subset using the same settings mentioned previously, but we included the filtering option to retain loci found in all four populations (-p 4). We used custom Python scripts to convert ddRAD fasta files to G-PhoCS format (Maier et al., 2019). We kept 569 loci present in all populations. We estimated demographic parameters under three different models: no migration, full migration (BR to CR and CR to BR), and BR to CR. The "no migration" model indicates that all shared alleles between BR and CR are due to incomplete lineage sorting, while the "full migration" and "BR to CR" indicate ongoing migration. All runs were performed using the same settings of 850000 MCMC steps, sampling every 100 generations with α = 1.0 and β = 1000.0 for the gamma distribution used for all priors of τ and θ parameters, and α = 1.0 and β = 0.00001 for the gamma distribution used for migration rates. The fine-tune option was set as "true," and values were the same as in Gronau et al. (2011). Convergence (ESS > 200) and posterior distributions were assessed in Tracer v1.6. Because we do not have estimates of genome mutation rate for crustaceans, we used the *Drosophila melanogaster* mutation rate (μ) of 3.8×10^{-6} per site per generation time (Lynch, 2010) to transform estimates of θ to N_e ($\theta = 4N_e \mu$), τ to divergence time ($\tau = T \mu$), and m_{SX} (mutation-scaled instantaneous migration from population S to population X) to migrants per generation ($msx \times \theta x/4 = Msx$) when migration was included in the model (Grounau et al., 2011).

It is not possible to compare different models using G-PhoCS, so we used a complementary coalescent approach using Migrate-N v.4.2.14 (Beerli & Felsenstein, 2001; Beerli, 2006) to compare different scenarios. The same custom python scripts to convert ddRAD fasta files to Migrate-N format were used (Maier et al., 2019), and we also used the same subset of individuals and loci analyzed on G-PhoCS. We estimated parameters under five different models: 1) full migration model; 2) migration

from CR to BR; 3) migration from BR to CR; 4) no migration model, and BR splitting from CR; 5) no migration model, and CR splitting from BR. We compare these models to understand if shared alleles among groups result from ongoing migration (models 1, 2, and 3) or ancestral polymorphism (models 4 and 5). Uniform priors were set for θ (0.000010 - 0.01) and M (0 – 1000) for all models, and split (0.000010 - 0.01) and split standard deviation (0.000010 - 0.01) for models 4 and 5, which included a splitting event. We conducted analyses using four heated chains (1.0, 1.2, 3.0, and 10000.0) ran for 20 000 steps, sampling every 100 generations, after 20 000 steps were discarded as burn-in.

Results

mtDNA (COI)

We used 66 sequences (570 bp), including extra sequences from North Carolina (considered as a "Caribbean" sequence in the haplotype network), North Brazil, and French Guiana (within the Amazon-Orinoco plume) not analyzed in Peres & Mantelatto (2020). The haplotype network depicts a clear split between the individuals from North and South of the Amazon-Orinoco plume, but one from the Caribbean fall within the "Brazilian" network. Individuals from the Amazon-Orinoco plume region show the most common haplotype shared by "Brazilian" individuals, though they are geographically closer to the Caribbean individuals (Figure 2).

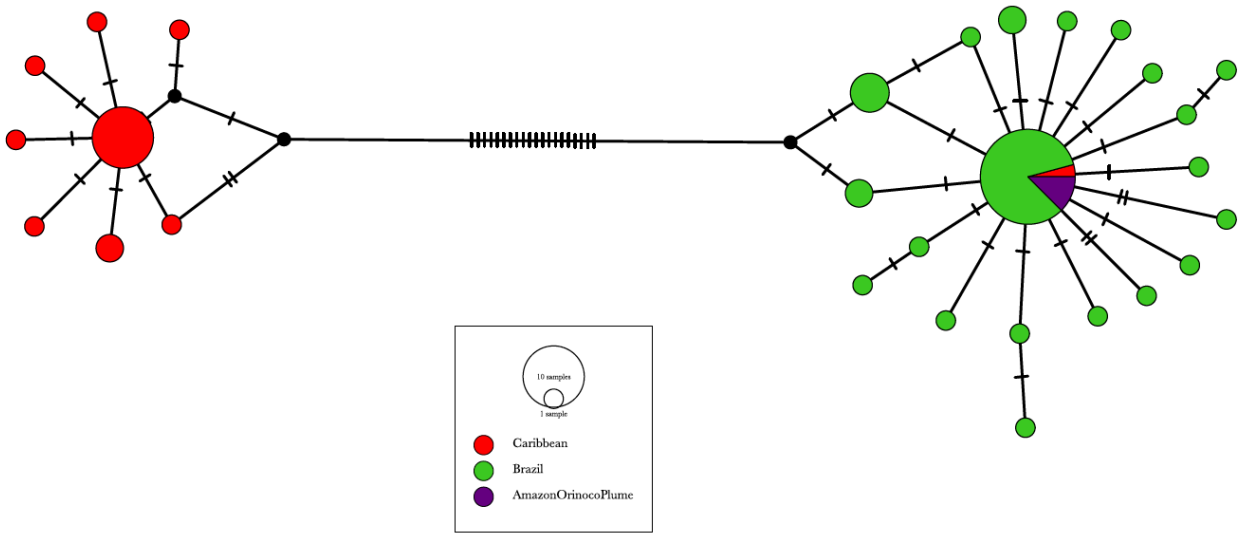


Figure 2. Haplotype network result for *Callinectes ornatus*. The size of the network circles is proportional to the haplotype frequency. Colors represent the region from where the haplotype was sampled.

ddRAD-seq

After process_radtags, we retained 570 million reads. The *denovo* pipeline ended up with ~300 thousand loci (mean = 201.2 bp) with effective per-sample coverage ranging from 15.4x to 469.9x. Of the 63 individuals, we excluded 18 that showed fewer reads as mentioned in the Methods section. The number of loci and SNPs for each dataset is shown in Table 1. The highest number of loci and SNPs was found in the dataset n = 30/-r 0.25 (42336 loci; 8202 SNPs), and the lowest number was found in the dataset n = 45/-r 0.75 (634 loci; 324 SNPs), after the removal of 24 non-neutral SNPs identified by both BayeScan and PCAdapt.

Outlier detection

After blasting all the 24 non-neutral loci, we found ten hits matching sequences deposited in public databases (Table 2). Out of these hits, two hits represented annotated genes: nascent polypeptide-associated complex subunit alpha and Down syndrome cell adhesion molecule isoform (Dscam) gene.

Table 2. Result of the blastn search using non-neutral loci identified by BayeScan and PCAdapt. Searches were restricted to Decapoda (organism:decapoda (taxid:6683)).

Description	Scientific Name	Max Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Macrobrachium nipponense isolate FS-2020 chromosome 16	*Macrobrachium nipponense*	47.3	24%	0.004	86.84%	92354735	CP062020.1
Macrobrachium nipponense isolate FS-2020 chromosome 17	*Macrobrachium nipponense*	42.8	25%	0.049	83.33%	89928525	CP062018.1
Macrobrachium nipponense isolate FS-2020 chromosome 21	*Macrobrachium nipponense*	48.2	36%	0.001	77.05%	84091324	CP062019.1
Macrobrachium nipponense isolate FS-2020 chromosome 36	*Macrobrachium nipponense*	52.7	24%	1.00E-04	85.71%	66999283	CP062024.1
Macrobrachium nipponense isolate FS-2020 chromosome 38	*Macrobrachium nipponense*	44.6	22%	0.016	85.37%	62717774	CP062056.1
Macrobrachium nipponense isolate FS-2020 chromosome 38	*Macrobrachium nipponense*	44.6	23%	0.014	85.37%	62717774	CP062056.1
Macrobrachium nipponense isolate FS-2020 chromosome 47	*Macrobrachium nipponense*	57.2	22%	3.00E-06	86.96%	36926747	CP062011.1
Penaeus monodon Down syndrome cell adhesion molecule isoform (Dscam) gene, partial cds; alternatively spliced	*Penaeus monodon*	179	97%	4.00E-43	78.85%	267116	MK838771.1
PREDICTED: Penaeus monodon nascent polypeptide-associated complex subunit alpha, muscle-specific form-like (LOC119568284), mRNA	*Penaeus monodon*	50.9	17%	5.00E-04	87.50%	2808	XM_037916742.1

Population structure

The pairwise-F_{ST} using neutral datasets indicates high and significant divergence between all populations from Brazil (named as BR from here) and the Caribbean group (called as CR from here), usually showing $F_{ST} > 0.40$. However, when using datasets showing <672 SNPs, most F_{ST} estimates were <0.30. In all datasets, divergence among BR populations was always lower than BR x CR comparisons and not significant (Table 3). Regarding pairwise-FST using non-neutral datasets, we found almost complete and significant differentiation ($F_{ST} > 0.98$) in all comparisons but one (SNPs = 372).

Table 3. Pairwise-F_{ST} results for all nine datasets. Lower diagonal: neutral loci; Upper diagonal: non-neutral loci. Bold values represent significant comparisons.

	n = 30				n = 40				n = 45				
-r 0.25		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**
	CR	-	**0.994**	**0.992**	**0.993**	**CR**	-	**0.997**	**0.995**	**0.993**	**CR**	-	0.986	0.992	0.988
	NE	**0.447**	-	**0.44**	**0.472**	**NE**	**0.42**	-	-0.123	0.078	**NE**	0.417	-	0.203	0.145
	SE	**0.431**	0.022	-	0.026	**SE**	**0.44**	0.005	-	0.065	**SE**	0.417	0.016	-	
	S	**0.435**	0.014	0.015	-	**S**	**0.417**	0.011	0.009	-	**S**	0.425	0.002	0.017	-
-r 0.50		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**
	CR	-	**0.998**	**0.995**	**0.996**	**CR**		**0.992**	**0.994**	**0.99**	**CR**	-	**0.999**	**0.997**	**0.998**
	NE	**0.466**	-	-0.026	-0.019	**NE**	**0.47**	-	-0.051	0.1	**NE**	**0.491**	-	-0.025	-0.051
	SE	**0.457**	0.011	-	-0.012	**SE**	**0.48**	0.009	-	0.06	**SE**	**0.473**	0.002	-	0.015
	S	**0.455**	0.013	0.009	-	**S**	**0.497**	0.007	-0.005	-	**S**	**0.529**	0.01	0.007	-
-r 0.75		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**		**CR**	**NE**	**SE**	**S**
	CR	-	**0.989**	**0.999**	**0.996**	**CR**	-	0	**0.992**	**0.991**	**CR**	-	0	0	0
	NE	**0.357**	-	0.055	-0.068	**NE**	**0.227**	-	0.135	-0.107	**NE**	**0.291**	-	-0.072	-0.046
	SE	**0.446**	0.012	-	0.011	**SE**	**0.319**	0.003	-	0.026	**SE**	**0.1**	0.043	-	0
	S	**0.501**	-0.01	0.005	-	**S**	**0.587**	0.002	0.027	-	**S**	0.192	-0.017	-0.031	-

STRUCTURE results for neutral and non-neutral loci are shown in figure 3. Depending on the dataset, we found different most probable K values. Overall, most results show a CR and BR genetic structure, but individuals are mixed. Like pairwise-FST, we did not find support for genetic structure among Brazilian populations. Datasets showing <672 SNPs could not differentiate any group. The adaptive divergence completely separates CR and BR.

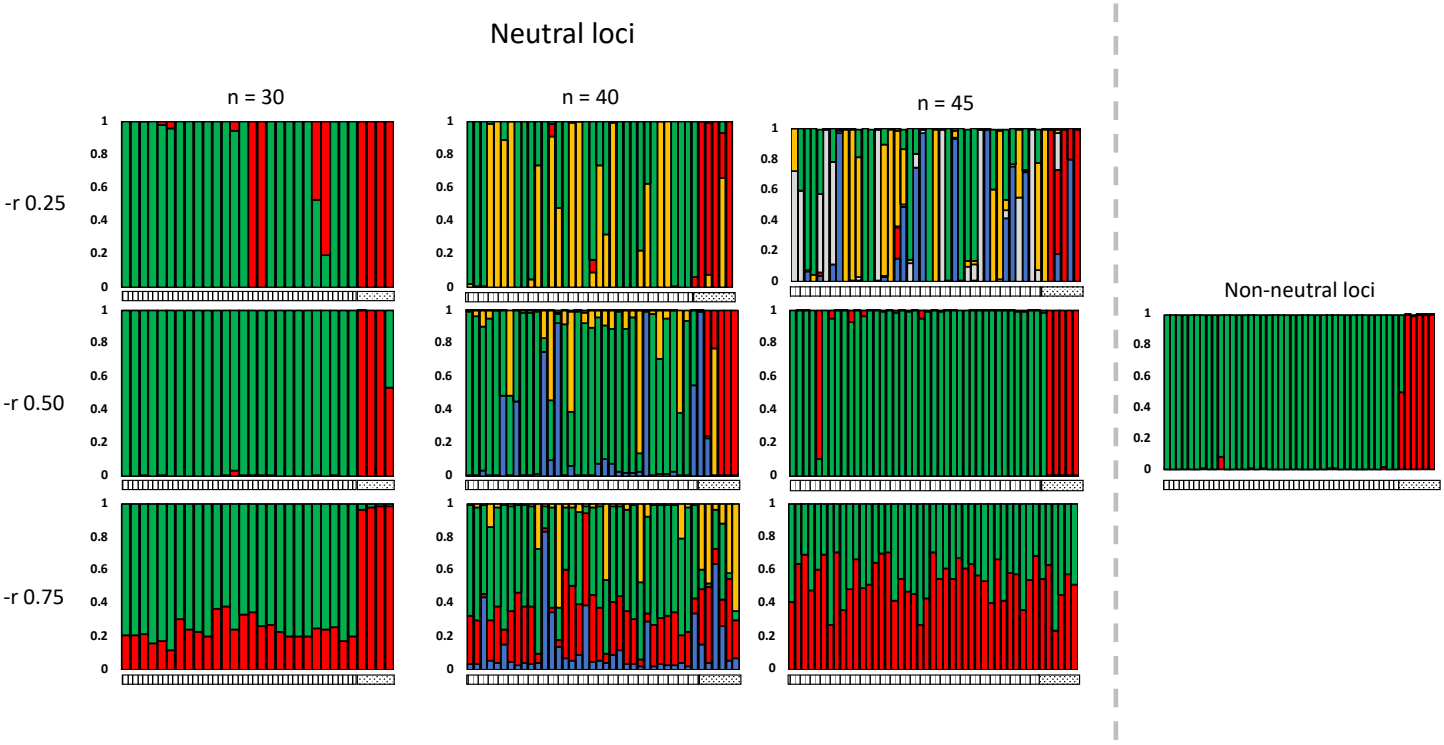


Figure 3. STRUCTURE plots for each dataset analyzed. Each vertical bar represents on individual. Different colors represent different genetic clusters (K). Horizonal bars below each plot represent Brazilian populations (vertical line texture) and Caribbean (dot texture) groups.

Divergence time, migration, and effective population size

The estimated divergence time between CR and BR from G-PhoCS analyses varied depending on the model tested. The "no migration" model indicates that CR and BR individuals separated 82.7KYA (23.6 – 149.6KYA), the "full migration" model 2.6 MYA (1.9 – 3.3 MYA), and the "BR to CR migration only" model 2.7MYA (2 – 3.1MYA). The migration rate from BR to CR is approximately 16 times higher than from CR to BR in the "full migration" resulting in asymmetrical migration. When we excluded the CR to BR migration band, the demographic parameters were almost not affected, reinforcing the low estimates of this migration route. Also, we detected a much greater population size in BR than in CR in all simulations (Figure 4).

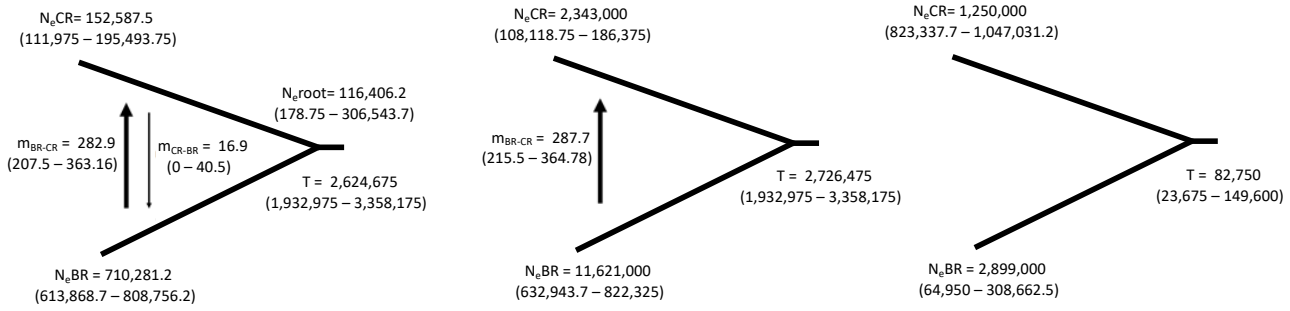


Figure 4. G-PhoCS models and estimated demographic parameters. From right to left: full migration, BR to CR, and no migration model. N_e: Effective population size; m = migrants per generation; T = divergence time in years.

The most supported Migrate-N model was the full migration model (1), which also showed an asymmetrical migration scenario with greater migration from BR to CR than from CR to BR (Table 3). This result indicates an ongoing migration scenario and that shared alleles are not caused by ancestral polymorphism.

Table 3. Migrate-N results of each compared model. Ln model: marginal Likelihood of the model; LBF: Log Bayes Factor.

Model	Ln model	LBF
full model	-154476.3	0
BR to CR	-154664.38	-188.08
CR to BR	-154695.84	-219.54
no migration model (BR splitting from CR)	-155594.22	-1117.92
no migration model (CR splitting from BR)	-155694.38	-1218.08

Discussion

Our results indicate a complex diversification scenario within *Callinectes ornatus*. We found that populations at the North and South regions of the Amazon-Orinoco plume show neutral and adaptive structure, and at the same time, show asymmetrical migration and mixture individuals at neutral loci. This result challenges the validation of *C. ornatus* as a single entity, but the two lineages we found also do not represent clear separated species, as we will discuss below. Therefore, employing mtDNA and ddRAD-seq, we show evidence of an incipient speciation scenario within a widespread species along the Western Atlantic, showing high dispersal potential. These results corroborate the assertion that the Tropical Atlantic is an exciting region to investigate marine diversification due to its intriguing lack of common patterns among species (Floeter et al., 2008). For instance, on the Western Atlantic, the Amazon-Orinoco plume may or may not act as a barrier depending on the species (Joyeux et al., 2001).

High-resolution markers can detect fine-scale structure when present in benthic marine species showing high dispersal potential due to their larvae (Benestan et al., 2015; Vendramini et al., 2017; Xuereb et al., 2018). However, our results show that there is no structure across populations along the Brazilian coast. Many marine and estuarine species show a lack of structure in the same region, such as annelids, fishes, hermits and crabs (Laurenzano et al., 2013; Nunes et al., 2017; Buranelli et al., 2019; Nishikawa et al., 2021). The novelty of our results is to show that not even SNPs could detect fine-scale structure along the Brazilian coast in *C. ornatus*, at least considering the populations we sampled. This pattern is probably influenced by the South Equatorial Current (SEC) (Lumpkin & Johnson, 2013). The SEC branches in the southward Brazil Current – BC and the northward North Brazil Current (NBC) (Silveira et al., 2000; Lumpkin & Johnson, 2013). The BC, its gyres, and local currents are probably the factors influencing the lack of structure in many populations along the Brazilian coast by transporting larvae (Silveira et al., 2000; Lumpkin & Johnson, 2013). Some mangrove trees show differentiation between the North Coast and East/South coast of Brazil (Mori et al., 2015). Notably, adding mtDNA data of samples from the North coast of Brazil and regions within the Amazon-Orinoco plume did not detect differentiation, even considering the potential of the effects of the split of the SEC. Unfortunately, we could not analyze these samples using ddRAD-seq, and this result might be constrained by the COI resolution, and further studies should explore this topic.

Contrastingly, we found great genetic differentiation between the BR and CR regions in all analyses and in both markers. Here, we added to Peres & Mantelatto (2020) COI dataset sequences from the northern coast of South America (including sequences from within the plume) and North Carolina samples. As expected, North Carolina samples fall within the CR network, but sequences from within the plume are represented by the most common haplotype in the BR network. This is surprising because they are geographically near the CR group and considering the NBC, which would favor transport northwards. Considering ddRAD-seq, pairwise-F_{ST} and STRUCTURE plots find a clear BR and CR

group in all datasets tested. However, the STRUCTURE plots also show that it is not a complete differentiation, which we would expect if we were dealing with more than one species (e.g., Hughes et al., 2019; Pertierra et al., 2019). We find individuals being assigned to the opposite cluster despite their location and individuals showing signs of mixture and mitonuclear discordance. For instance, an individual from Trinidad & Tobago has the CR mtDNA but is assigned to the BR group, while a Saint Martin individual is assigned to the BR group based on mtDNA but is assigned to the CR cluster under the ddRAD analysis.

We modeled gene flow to answer if what we found resulted from migration between populations or incomplete lineage sorting. The latter is based on the fact that there are sister species separated by the Amazon-Orinoco plume (Rocha, 2003; Trovant et al., 2016). Thus, our scenario could be explained by complete isolation but not complete divergence due to large effective population sizes reducing genetic drift (Allendorf et al., 2010). Supporting this hypothesis, a time-calibrated mtDNA phylogeny estimated the divergence of both groups around 4.6 MYA (2.7–6.5 Mya), consistent with the complete establishment of the plume (Peres & Mantelatto, 2020). Under an allopatric scenario (no migration), we estimated a divergence dating back to 82.7 KYA (23.6 – 149.6 KYA), not matching the barrier age. However, our most probable model is an asymmetrical migration scenario, being migration from BR to CR higher than in the opposite direction. Under an isolation-with-migration model, we found a divergence time of 2.6 MYA (1.9 – 3.3 MYA). We believe this estimation is more robust based on the number of independent loci used and less variance of the estimated date than the previous dating using mtDNA, which may not reflect the species tree (Ballard & Whitlock, 2004). Interestingly, the estimated date using ddRAD loci agrees precisely with a further Andean uplift and Quaternary glaciation resulting in increases in freshwater and sediments outflow, which might have impacted marine species in that region (Figueiredo et al., 2009; Hoorn et al., 2010). At the same time, the NBC was originated due to changes in Atlantic currents caused by the closure of the Isthmus of Panama (Heinrich & Zonneveld, 2013). We estimated a migration rate of 282 migrants per generation in the BR-CR direction and 16 in the CR-BR direction, which agrees with the flow of the NBC. Using the same approach (G-PhoCS),

Bertola et al. (2020) found much lower estimates (0.02 – 1.59) yet no genetic structure between fish populations from the Gulf of Mexico and Florida Atlantic coast. Although this can be justified by their low estimates of effective population size, it is intriguing why we found high rates of migration rates and strong genetic differentiation.

A possible explanation for this pattern is that *C. ornatus* is diversifying into two different entities. Speciation can be thought of as a continuous process that eventually results in complete reproductive isolation (Abbott et al., 2013). In marine systems, speciation with gene flow is highly likely (Miglietta et al., 2011). Under a general model of speciation with gene flow, the first phase represents positive selection on a few genes while most genome shows low differentiation (Feder et al., 2012). Later, divergence hitchhiking creates "islands of differentiation," and, posteriorly, genome hitchhiking led to great differentiation across the whole genome (Feder et al., 2012; Martin et al., 2013). Therefore, traits or genes can show divergence in the absence of evolutionary independence when gene flow is prevalent (Hey & Pinho, 2012). Empirical results also confirm this possibility, as in the eels *Anguilla anguilla* and *A. rostrata* that show variation in gene flow rate across the genome, indicating some regions are more resistant to introgression despite high gene flow between species (Nikolic et al., 2019). When we look at putative loci under selection (non-neutral loci), we found that BR and CR are entirely differentiated. This is probably not a false-positive result because we used more than one method to identify such loci (Narum & Hess, 2011; Villemereuil et al., 2014; François et al., 2016). Unfortunately, we could not match them with annotated genes or genes we could use to interpret a function differentiating regions. We may hypothesize that BR and CR populations are accumulating neutral and adaptive differentiation despite gene flow. This is also supported by the fact that migration seems to be primary in one direction, implying that BR is independent of CR, and considering that the CR effective population size is smaller than BR, indicating a faster rate of genetic drift (Hey & Pinho, 2012). Disruptive selection can be contributing to our results, but also neutral differences linked to essential traits and both impacting mating (prezygotic isolation) or hybrids viability (postzygotic isolation) (Coyne & Orr, 1998).

Callinectes ornatus abundance and reproduction, for example, seems to be affected by environmental conditions (Haefner, 1990; Mantelatto & Fransozo, 1999; Mantelatto 2000). If a high number of migrants (larvae) reaches new populations that are under a different seasonal pattern, we might have a mismatch between the south and north hemisphere individuals mating (Lowerre-Barbieri et al., 2011). However, we are inclined to affirm that prezygotic isolation is not the primary factor affecting *C. ornatus* diversification because we found potential hybrids (individuals showing mtDNA from one group and nuDNA from the other). One of the most probable hybrid identified due to mitonuclear discordance across all datasets is from Trinidad & Tobago, right at the north part out of the plume. We also found many mixed individuals in which a clear assignment to a genetic cluster was not possible. Although we did not get ddRAD data from specimens within the plume, we hypothesize this region is a potential hybrid zone. The environmental changes caused by the Amazon-Orinoco plume can lead to local adaptation in *C. ornatus* from this area. Local adaptation can generate localized genome changes and combined with directional migration, might lead to genome instability in hybrids or assortative mating, increasing differentiation (Dion-Côté & Barbash, 2017; Kulmuni et al., 2020). Genes under strong positive divergence are more prone to show Dobzhansky–Muller incompatibilities, supporting our hypothesis (Orr & Turelli, 2001).

Callinectes ornatus shows low tolerance for salinity variance, so the intensification of freshwater and sediment discharge at the final stages of the plume formation might have started the divergence between CR and BR (Norse, 1978; Garçon et al., 2007). Thus, our hypothesized scenario proposes a combination of an environmental barrier and disruptive selection to explain the genetic divergence with high gene flow (Potkamp & Fransen, 2019). Although we could not detect the involved genes to determine if hybrids are under postzygotic isolation and we did not test assortative mating, similar scenarios occur in other high dispersal species. The *Rhagada* land snail shows two different habitat-related morphology and an intermediate morphology in a narrow habitat transition (Stankowski, 2013). The authors argue that despite gene flow, disruptive selection is acting upon the morphologies alongside postzygotic selection, disfavoring the hybrid morphology, suggesting speciation. Also, the Atlantic cod *Gadus*

morhua show divergence in genes involved in oxygen use and osmoregulation, which are probably located in huge blocks of chromosomic inversions, generating local adaption and adaptive and neutral structure among populations (Barth et al., 2017). As an operational but debatable criterion, $F_{ST} > 0.35$ can be considered a threshold to define two separate species under an isolation-with-migration model (Hey & Pinho, 2012). Therefore, it is feasible to affirm we captured an incipient speciation scenario.

ddRAD-seq and mtDNA comparison

For a long time, the mtDNA has been used to investigate the genetic structure, phylogeographic patterns, and systematics due to many of its properties (Moritz et al., 1987). The results shown by this marker might not always agree with nuclear DNA, an event called mitonuclear discordance (Toews & Brelsford, 2012). There are examples of mtDNA not capturing structure found by SNPs (Pedraza-Marrón et al., 2019), so as cases in which mtDNA delineates several cryptic species that are not supported by SNPs data (Hinosoka et al., 2019). Because mitochondrial and nuclear genomes are under different evolutionary processes, several mechanisms might explain the discordance like incomplete lineage sorting, sex-biased dispersal, hybridization, and selective sweeps (Ballard & Whitlock, 2004; Galtier et al., 2009; Toews & Brelsford, 2012; Edwards et al., 2016). In opposition, our results show an agreement between both types of markers. An advantage of using markers with different types of inheritance is the possibility of finding signs of introgression and hybridization by comparing these markers (Arnold, 1992). A previous study has shown 4% COI divergence between the *C. ornatus* CR and BR group, a value that is under the interspecific 6% gap for the genus but still high (Lefébure et al., 2006; Silva et al., 2011), and also found a single haplotype shared by both regions (Peres & Mantelatto, 2020). We expected to see two well-separated species or mitonuclear discordance (i.e., lack of nuDNA differentiation). But, in fact, we show individuals admixture caused by the combination of disruptive selection and ongoing asymmetrical migration, and not incomplete lineage sorting. Additionally, we identified individuals with mtDNA genome from one group and nuDNA from the other, reinforcing our

hypothesis of incipient speciation in which reproductive isolation has not taken place. Sometimes, introgression is found in just one genome (nuDNA: Beheregaray et al., 2017; mtDNA: Hughes et al., 2019), but the factors just causing one or another are still debated (Toews & Brelsford, 2012). We believe *C. ornatus* results of a particular combination of large population sizes, recent divergence, migration, and disruptive selection, as we discussed. Overall, although markers resulted in the same pattern, ddRAD-seq offers a better resolution of the diversification process. Future studies could take advantage of delineating experiments based on previous mtDNA results or consider employing both markers to investigate possible complicated diversification scenarios.

ddRAD-seq, missing data, and the use of multiple datasets

An intrinsic feature of ddRAD-seq is the occurrence of missing data (Andrews et al., 2016). Because of how the technique works, it is recommended to use high molecular weight DNA (1000 ng of DNA), more feasible to be found in freshly collected samples, even though it can be used for small yields (100ng) of DNA (Peterson et al., 2012; Puritz et al., 2014). However, many attempts have succeeded using museum specimens or individuals under different degradation levels in ddRAD-seq approaches (Graham et al., 2015; Battey & Klicka, 2017; Haponski et al., 2017). This shows that there is a tolerable level regarding the DNA quality that does not restrain the use of ddRAD-seq in such kinds of specimens. This opens a window to explore the vast diversity of museum specimens collected in different regions and times, especially species not easy to collect like some marine animals, and still use ddRAD-seq to answer an array of questions, especially the ones in the intraspecific or shallow levels. However, artifacts may emerge when using museum specimens as we have used. This led us to exclude some individuals from our analyses and test a combination of different datasets.

Using such strategy, we show that all datasets indicate similar results despite the number of individuals or SNPs excluded. The effect of excluding missing data has been shown to impact analyses by reducing phylogenetic inference accuracy (Huang & Knowles, 2016). On shallow scales, the use of

larger datasets is preferable despite showing missing data, being able to resolve relationships even with 90% of missing data (Tripp et al., 2017; Bombonato et al., 2020). On the populational scale, genetic diversity and differentiation seem not to be highly affected by missing data when a large number of SNPs are being analyzed (Fu, 2014). Additionally, small sample sizes as low as two individuals per population also do not influence population differentiation estimates if analyzing >1000 SNPs (Willing et al., 2012; Nazareno et al., 2017). We acknowledge the limitations of our dataset and add empirical results to this discussion showing that excluding missing data and individuals did not change overall outcomes. We show changes in the best K estimated by the STRUCTURE analysis and slight changes in F_{ST} estimates that do not compromise our interpretation. However, using <600 SNPs hindered population differentiation as seen on STRUCTURE plots, yet F_{ST} was still high (around 0.3), but smaller than other datasets. Due to divergence with gene flow, it is expected a scenario of few regions of the genome showing great differentiation ("islands of differentiation") against a homogenous and not so differentiated background (Feder et al., 2012). Whole-genome sequencing of species showing a complex process of diversification indeed shows this exact pattern (Martin et al., 2013). Considering we explored multiple datasets, we feel confident affirming *C. ornatus* is going through an incipient speciation scenario and suggest it as an approach when facing missing data to avoid wrong or limited conclusions (Díaz-Arce & Rodríguez-Ezpeleta, 2019).

Conclusions

The patterns and processes of marine species diversification still have a lot to be uncovered (Miglietta et al., 2011; Bowen et al., 2014). Our work contributes to the field by employing mtDNA and ddRAD-seq to investigate the diversification within the widespread *C. ornatus* along the tropical Western Atlantic. We show that this species is under a divergence with gene flow process, characterizing an incipient speciation scenario. Disruptive selection acting upon individuals from the BR and CR region coupled with a permeable barrier (the Amazon-Orinoco plume), differences in effective population size,

and oceanographic currents may be the main factors in this complex scenario. Further studies should get more samples from the potential hybrid/transition zone between the CR and BR groups to better understand the population-species continuum (Losos & Glor, 2003; Edwards et al., 2016). Until this moment, no morphological character was found that could separate the species so as no discussion that *C. ornatus* might actually represent two different entities (Williams, 1974; Santos 2007; personal communication). We take a conservative approach to not nominate a new species because there are no operational criteria to empirically separate the species (de Queiroz, 2007). However, we believe we captured an early moment in the speciation process and can affirm we are dealing with two separately evolving metapopulation lineages (de Queiroz, 2007).

CHAPTER 3 (Manuscript *in prep.*)

Genetic diversity spatial trends in marine crabs: interspecific latitudinal gradient and species idiosyncratic patterns.

Pedro A. Peres & Fernando L. Mantelatto.

Introduction

Explaining the patterns and process of biodiversity is one of the main challenges of ecology and evolution. Within this goal, questions addressing the genetic diversity (GD) spatial patterns are still open. Species richness tends to be higher towards the tropics on a global scale, a spatial pattern known as the Latitudinal Diversity Gradient (LDG), one of the most intriguing and well documented biological patterns (Hawkins, 2001; Hillebrand, 2004a; Kinlock et al., 2017). Although there is no consensus on the determinants of this pattern, most explanations fall within historical, biogeographical, and/or ecological processes (Mittelbach et al., 2007; Lawrance & Fraser, 2020). Recent studies addressing if the LDG extends to the GD found support for a broad-scale Latitudinal Genetic Diversity Gradient (LGDG) (Miraldo et al., 2016; Gratton et al., 2017; Schluter & Pennel, 2017). Similar to the LDG, the LGDG shows a spatial pattern of higher GD at lower latitudes and decreasing GD towards higher latitudes. However, the LGDG remains poorly explored, and there are unresolved questions about whether this is a general trend at the intra- and/or interspecific levels across different taxa.

The LGDG has been mostly explored at the species level, focusing on intraspecific GD of mainly terrestrial vertebrates such as mammals, birds, and amphibians (e.g., Adams & Hadly, 2013; Miraldo et al., 2016), and less presently on invertebrates (e.g., Schär et al., 2016). Despite the taxa, the LGDG is often found, indicating a possible common cause for all cases. A potential explanation for the LGDG is the Evolutionary Speed Hypothesis (ESH – Rohde, 1992). The ESH states that individuals at lower latitudes show higher metabolic rates, shorter generation times, and higher mutation rates caused by temperature (Rohde, 1992). The consequences of this climate-based hypothesis are higher diversification rates within and among species at lower latitudes (Mittelbach et al., 2007). As intra- and interspecific GD may be governed by the same forces (Antonovics, 2003; Vellend et al., 2014), the ESH offers a

plausible explanation for the latitudinal gradient patterns over different levels, even though other factors might also be contributing (Gillman & Wright, 2014). Currently, advances in molecular genetics have strengthened the ESH as GD data has become available for a great range of species (Mittelbach et al., 2007). For instance, combining life cycle experiments and GD assessments confirmed the ESH as the driver of LGDG in a non-biting midge species (Oppold et al., 2016). Also, factors related to latitude are determinants for substitution rates in turtles (Lourenço et al., 2013), and bumblebees from warmer regions show longer branch lengths than other areas (Lin et al., 2019).

Although highly prevalent, the LGDG might not always explain the GD spatial distribution. An alternative to latitude-related hypotheses for the intraspecific level is the Central-Marginal Hypothesis (CMH) (Eckert et al., 2008). Instead of latitudinal trends, GD would be spatially distributed according to the species distributional range. Higher GD is found at the core of the distribution while declining towards peripheral populations. Following the abundance center model, species tend to show higher abundance at the core of their distribution, where they find optimal conditions and expand their range to sub-optimal peripheral areas showing smaller abundance (Sagarin & Gaines, 2002). This model has different consequences on effective population size and migration rates among central and marginal populations, directly affecting spatial GD distribution and generating CMH patterns in many taxonomic groups (Eckert et al., 2008). Interestingly, the CMH and LGDG can be combined, resulting in higher GD at the central areas of the distribution than at the marginal populations. Simultaneously, populations on the lower latitudes side show higher GD than higher-latitude populations (Guo, 2012). Variations within this interaction can be found depending on the species range orientations, such as a more North-South or West-East distribution and are anticipated by the CMH-LGDG hypothesis (Guo, 2012). Either way, many species show a better adjustment between GD and latitude under non-linear models (CMH-LGDG) than linear models (LGDG) (Guo, 2012). It is essential to notice that most of the studies testing CMH-LGDG and/or LGDG expect non-linear and linear relationships respectively as they deal with species ranging on just one hemisphere. However, LGDG might be depicted by non-linear models when

the species range encompasses both hemispheres, as shown in interspecific patterns (e.g., Miraldo et al., 2016). Therefore, caution is needed to disentangle which hypothesis best explains each case.

As previously mentioned, most of the studies on the GD spatial distribution deal with terrestrial vertebrates or invertebrates from temperate regions (Eckert et al., 2008). There is a significant gap in knowledge concerning the patterns and processes of GD distribution across marine species (but see Manel et al., 2020 for a discussion on fishes), mainly tropical marine invertebrates (but see Liggins et al., 2014 for a discussion on echinoderms). The LDG is controversial when considering the marine environment. We find varying support in favor of the LDG depending on the taxa and region (Hillebrand, 2004b; Tittensor et al., 2010), but there are also cases showing the opposite response – a latitudinal inverse gradient (Rivadeneira et al., 2011). Furthermore, there is an indication of a bimodal latitudinal species gradient in marine environments, which shows higher diversity at intermediate latitudes (Chaudhary et al., 2016; Chaudhary et al., 2017), not following LDG. If GD follows the species richness gradient, we may expect finding different marine environment patterns than those described for terrestrial environments. Adding uncertainty to the issue, marine species might not show any GD spatial distribution trend due to the probability of connectivity among species in this complex system. The Marine species Hypothesis (MH) states that common features in marine invertebrates might hamper the emergence of GD spatial trends (Liggins et al., 2015). Features like large population sizes and high migration rates among populations caused by pelagic larvae would prevent specific population dynamics that could generate some of the spatial patterns described (Palumbi, 1994). For instance, differences among central and marginal populations (CMH) are not likely to occur, or alleles might be spread across the whole distribution range and not forming high GD regions (LGDG) if all populations are highly connected.

Here, we addressed the question of which hypothesis (LGDG, CMH, CMH-LGDG, MH, or other – Figure 1) better explain the GD spatial distribution trends at the intra- and interspecific level in the marine environment using western Atlantic marine tropical crabs as models. The western Atlantic is ideal for discussing latitudinal gradients due to its north-south arrangement, and its coastal environment

does not show extreme longitudinal variation. Crabs are a diverse, abundant, and a speciose group along the coast and are found in many habitats (Ng et al., 2008; Davie et al., 2015; Mantelatto et al., 2020), offering an opportunity to explore general patterns for this particular taxon. We selected 14 crab species based on their latitudinal distribution encompassing representatives whose ranges include North and South hemispheres (n = 12), or just South hemisphere (n = 2) and different habitats. Although there are no global assessments of crabs' diversity, crabs show a LDG in species richness within temperate western Atlantic and southeastern Pacific regions (South America coast) (Astorga et al., 2003; Fernández et al., 2009; Pappalardo & Fernández, 2014), and some mangrove crab families show LDG on a global scale (Sharifian et al., 2020). Thus, we expected to find an interspecific LGDG relationship due to previous information on the occurrence of species richness LDG, but also due to LGDG in terrestrial taxa (e.g., Miraldo et al., 2016). However, we hypothesize intraspecific GD spatial trends will vary across species. Depending on the species, a different hypothesis will be the best explanation. Additionally, we also tested if the ESH can explain intra- and interspecific GD spatial trends by assessing substitution rates at different latitudes. We expected to find longer branch lengths at lower latitudes.

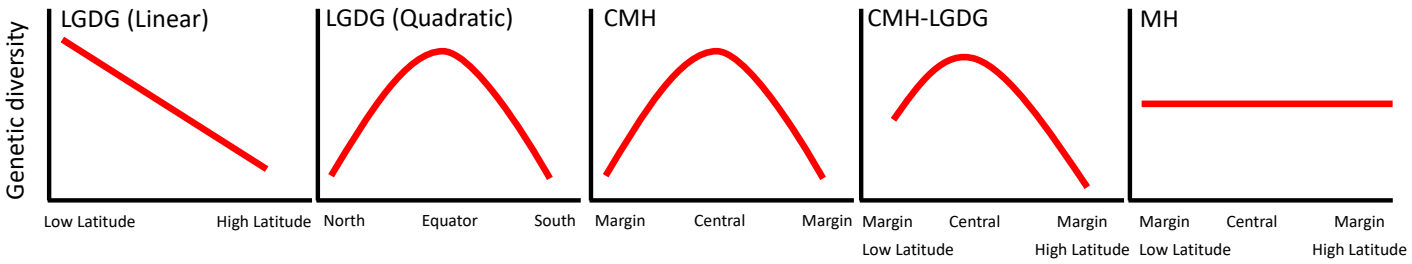


Figure 1. Representation of different models explaining genetic diversity spatial patterns.

Materials and Methods

Crab species

We studied *Acanthonyx petiverii* H. Milne Edwards 1834 of the Epialtidae; *Eriphia gonagra* (J.C. Fabricius, 1781) of the Eriphiidae; *Goniopsis cruentata* (Latreille, 1803) of the Grapsidae; *Leptuca leptodactyla* (Rathbun, in Rankin, 1898), *Leptuca thayeri* (Rathbun, 1900), *Minuca rapax* (Smith, 1870), and *Ucides cordatus* (Linnaeus, 1763) of the Ocypodidae; *Arenaeus cribrarius* (Lamarck, 1818), *Callinectes danae* Smith, 1869, and *Callinectes ornatus* Ordway 1863 of the Portunidae; *Aratus pisonii* (H. Milne Edwards, 1837), *Armases angustipes* (Dana, 1852), and *Sesarma rectum* Randall, 1840 of the Sesarmidae; *Neohelice granulata* (Dana, 1851) of the Varunidae. These species inhabit rocky shores, or coastal infralittoral, or mangroves. *Armases angustipes* and *N. granulata* are restricted to South America while the other species occur along the Brazilian Coast and Caribbean.

Sequence data

The cytochrome oxidase I (COI) gene is often utilized for phylogenetic and phylogeographic studies across taxa, including crabs, and is universally available in public databases (Avise, 2000; Porter & Hajibabaei, 2018). Also, the COI was the marker of choice in studies on latitudinal trends (e.g., Adams & Hadly, 2013; Miraldo et al., 2016; Manel et al., 2020). Considering its availability and previous use, COI represents an ideal marker for a single-marker study on GD spatial patterns. A total of 800 COI mitochondrial DNA sequences belonging to the 14 species were used, encompassing their whole distributional range. Sequences were either directly obtained for this study or from publicly available sequences from previous studies (Ituarte et al., 2012; Laurenzano et al., 2016; Marochi et al., 2017; Tamburus & Mantelatto, 2016; Buranelli & Mantelatto, 2017, 2019; Zupolini et al., 2017; Buranelli et al., 2019; Thurman et al., 2019; Peres & Mantelatto, 2020). In the latter case, DNA extraction, PCR

conditions, editing, and alignment followed Peres & Mantelatto (2020) and then were submitted to GenBank. The sampling location of sequences was obtained on the published paper or by contacting the authors. We only used sequences from locations of $n \geq 3$ and discarded others to minimize the impact of sample size on the genetic diversity estimation and maximize the number of sampling locations used (Goodall-Copestake et al., 2012; Schär et al., 2017). All sequences were visually inspected in the alignment and short or low-quality sequences were note used. For each sampling location, we considered only absolute latitude values and pooled locations within the same latitudinal band (one degree) for the subsequent analyses (i.e., if one sampling location was at 23°34'45''S and the other at 23°12'56''S, both were considered from the 23° latitudinal band). Sampled locations ranged from 36°N to 32°S.

Intraspecific analyses

To access intrapopulation genetic diversity, we estimated nucleotide diversity (π) and haplotype diversity (h) for each latitudinal band for each species using DNASP (Rozas et al., 2017). Due to the relationship between π and h (Godall-Copestake et al., 2012), we performed a Principal Component Analysis (PCA) to reduce the dimensionality of this set of variables, combined both to get a single metric to represent genetic diversity – the Principal Component 1 (PC1). The PC1 was retrieved because it is the component that explains most of the variance within the dataset (>85% of the variation in all cases). As mentioned in the Introduction, GD spatial distribution along latitudinal bands may be explained by different models. To test which model better defines the relationship between genetic diversity (represented by PC1) and latitude, we used a hierarchical regression model approach comparing null, first, second, third, and fourth-order polynomial regressions. We used latitude as the predictor variable and PC1 as the response variable. The predictor variable was centered before inputting it into the model to guarantee the independence of the predictor variables' terms in polynomial regressions (Schielzeth, 2010).

To test if the substitution rate shows latitudinal trends (ESH - Rohde, 1992), we calculated the branch length (BL) from the ancestral node to the tips for each individual within each species. We maintained only one copy of each haplotype unless it was sampled at different latitudinal bands because identical haplotypes have the same BL. Thus, we kept identical haplotypes for the analysis when they were from different latitudinal bands. Otherwise, the haplotype was removed. This strategy was performed considering our interest in the relationship between BL and latitude. Although we kept identical haplotypes, they were associated with different latitudes. Then, we built a maximum-likelihood phylogenetic tree for each species on IQ-Tree (Kalyaanamoorthy et al., 2017; Minh et al., 2020). Branch support was estimated using ultrafast bootstrapping with 1000 replicates. We used *Homologenus malayensis* (NC026080) as an external group for all constructed trees because it is a sister species to all Eubrachyura crabs (as all other species analyzed) and because its mitochondrial genome is available, allowing us to align its COI sequence to all of our species completely. BL were calculated from the resulting phylogeny using the packages *ape* (Paradis et al., 2019) and *adephylo* (Jombart and Dray2010).

Interspecific analyses

After independently calculating π and h for each species per latitudinal band, we calculated interspecific GD per latitudinal band by averaging π and h across all species present (Miraldo et al., 2016; Manel et al., 2020). Latitudinal bands represented by just one species were excluded. Then, we did the same procedures described for intraspecific analyses and estimated PC1 and tested its association with latitude using hierarchical regression models. Accordingly, we also calculated the average BL per latitudinal band and tested its association with latitude.

All intra- and interspecific analyses were run in R ver.4.0.2 (R Core Team, 2020).

Results

Interspecific analyses

A bell-shaped relationship (quadratic association) explains GD and latitude relationship as a general trend when combining all species (Figure 2, Table 1). Higher GD is found towards the Equator, followed by GD decrease on both hemispheres. There is no trend regarding the relationship between BL and latitude when combining all species, indicating that ESH is not the process driving the pattern.

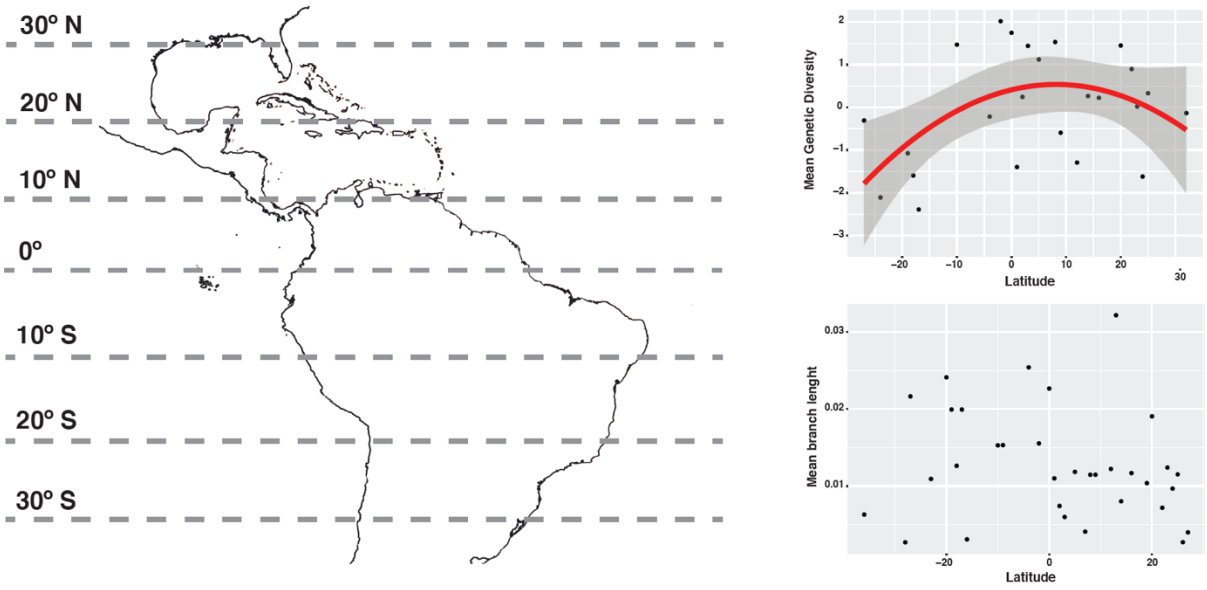


Figure 2. Map showing the latitude range analyzed and graphs of the interspecific results. The upper graphic depicts the significant association between latitude and mean genetic diversity (represented by PC1). The bottom graphic depicts the relationship between latitude and mean branch length (not significant).

Intraspecific analyses

The relationship between GD and latitude was not common across species (Figure 3, Table 1). The fourth-order polynomial model better explains the association between GD and latitude in *Aratus*

pisonii, *Leptuca thayeri*, and *Minuca rapax.* However, the pattern was not the same. There is a trend to higher GD towards low latitudes for *A. pisonii,* followed by a GD decrease at intermediate latitudes and an increase at the edge of the species range on both North and South hemispheres. *Minuca rapax* shows higher GD at low latitudes, a soft decrease on the North hemisphere latitudes but a steep decrease in the South hemisphere. The opposite happens for *L. thayeri*, which shows lower GD at low latitudes and higher GD at intermediate latitudes on North and South hemispheres. The second-order polynomial model better explains *Neohelice granulata* GD and latitude relationship. In the latter case, high GD is found at the middle of its distribution, followed by a decrease towards the edge of its distribution. Alternatively, all other species do not show any association between GD and latitude.

Table 1. Hierarchical Linear Regression results for the species showing significant association between GD/BL and latitude.

	GD					BL		
Species	Association with latitude	Model	adjusted-R2	F	p-value	Association with latitude	Model	adjusted-R2
Leptuca leptodactyla	NO	-	-	-	-	YES	fourth order	0.0219
Leptuca thayeri	YES	fourth order	0.1864	5.6609	0.0446	YES	fourth order	0.9564
Minuca rapax	YES	third order	0.6102	5.7106	0.43877	YES	first order	0.2709
Callinectes ornatus	NO	-	-	-	-	YES	fourth order	0.9862
Aratus pisonii	YES	fourth order	0.9812	232.455	<0.001	YES	fourth order	0.8103
Neohelice granulata	YES	second order	0.8941	22.4992	0.04169	YES	first order	0.1746
All species	YES	second order	0.1843	3.712	0.04803	NO		

The relationship between BL and latitude was also not common across species (Figure 3, Table 1). The fourth-order polynomial model better explains the association between BL and latitude in *A. pisonii*, *L. thayeri, L. leptodactyla,* and *C. ornatus* but depict different patterns. *Aratus pisonii'*s BL shows the same GD pattern found for this species, which is the same trend in *L. leptodactyla*. *Leptuca thayeri* and *C. ornatus* show higher BL in the North Hemisphere and lower BL in South Hemisphere. A linear model better explains BL and latitude relationship in *M. rapax* and *N. granulata.* All other species do not show any association between BL and latitude.

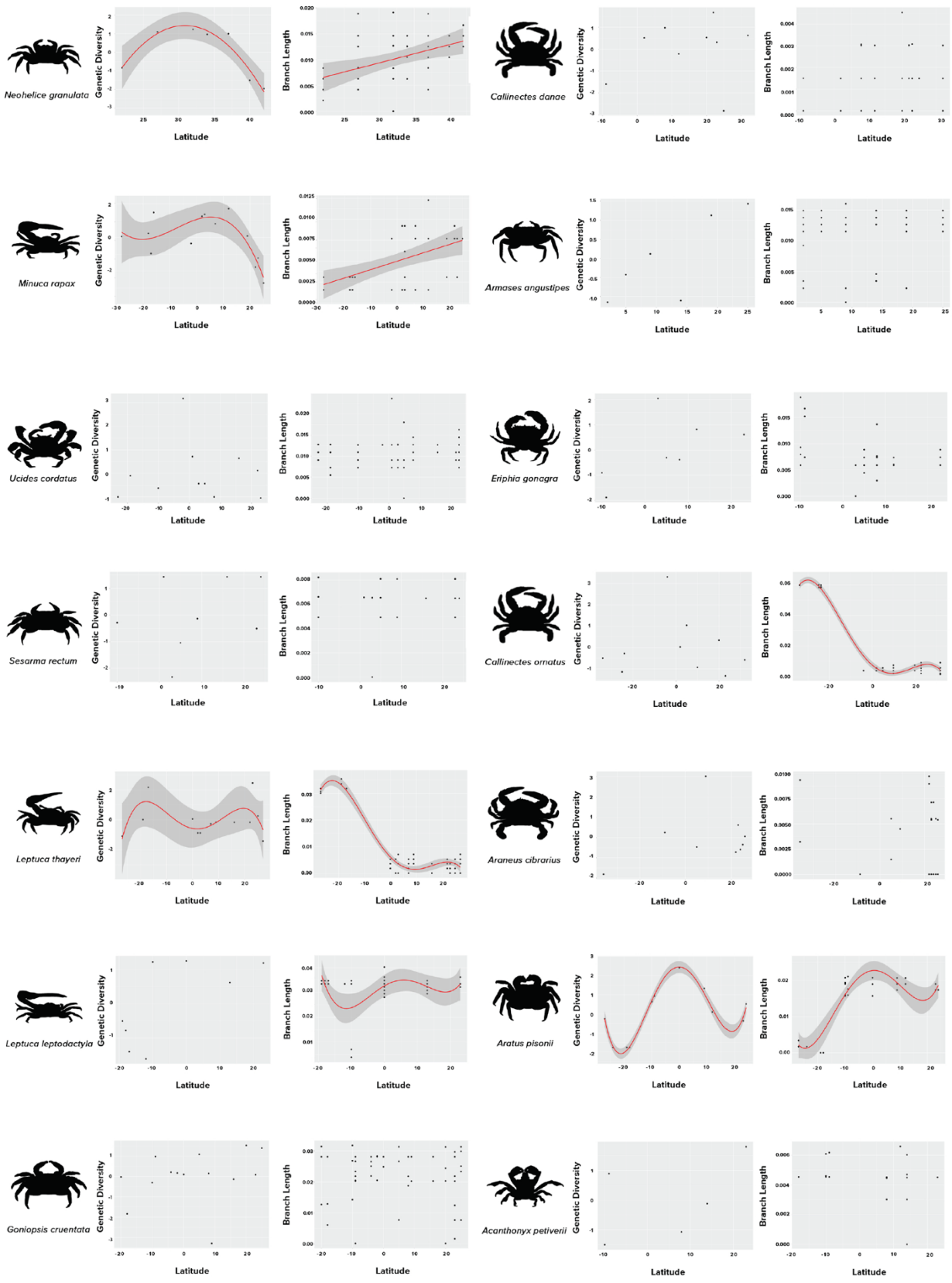


Figure 3. Results of the intraspecific comparison between latitude and genetic diversity of branch length. Only in significant associations a trend-line is included.

Discussion

Interspecific Latitudinal Genetic Diversity Gradient (LDGD) across crab species

Targeting co-distributed species of tropical western Atlantic marine crabs, we showed a LGDG at the interspecific level. Within latitudes 36°N to 32°S, GD depicts a bell-shaped (quadratic model) pattern peaking at lower latitudes and decreasing towards North and South hemispheres. LGDG has been found on terrestrial and marine vertebrates (Hillebrand, 2004b; Adams & Hadly, 2013; Miraldo et al., 2016; Manel et al., 2020), and for the first time, we confirm this pattern across marine crabs. Coastal marine species tend to show LGD (Tittensor et al., 2010), and works focusing solely on decapods or crabs also show this pattern on the limited scales they explored (Astorga et al., 2003; Fernández et al., 2009; Pappalardo & Fernández, 2014; Sharifian et al., 2020). Our results based on a wide range, reveal congruence between GD and crabs' species richness, indicating similar mechanisms underlying these patterns. Temperature is frequently stated as the most important variable explaining latitudinal trends (Astorga et al., 2003; Titterson et al., 2010; Manel et al., 2020). A common evoked hypothesis to explain species richness is the ESH (Rohde, 1992). The ESH states that temperature influences substitution rates leading to higher diversification on the tropics (Mittelbach et al., 2006; Dowle et al., 2013), and by extension, should also explain LGDG. We tested this hypothesis by assessing average BL at different latitudes, as we did with GD, expecting higher BL at lower latitudes. However, we did not meet our predictions, which indicates that substitution rates are not accelerated in this region, at least considering the molecular marker used.

The tropics are considered a cradle from where biodiversity originates (Chown & Gaston, 2000). Fossil evidence confirms that many marine benthic invertebrates' orders first appeared in the geological record in the tropics and expand to other regions (Jablonski, 1993; Jablonski et al., 2006; Kiessling et al., 2010). However, it is still not resolved from a molecular perspective if higher substitution rates are linked to the tropics being a center of biodiversity and how or if molecular evolution is influenced by temperature (Dowle et al., 2013). Some taxa show higher substitution rates in species from warmer regions (Wright et al., 2011). However, this does not seem a common pattern across the tree of life

(Jansson et al., 2013). Analyzing COI data of invertebrates and vertebrates, Orton et al. (2019) found that only around a little less than half of their dataset show higher substitution rates at lower latitudes indicating that the ESH is not as pervasive as assumed beforehand. Marine fishes show LDG, but higher speciation rates do not explain the species richness pattern (Rabosky et al., 2018). For fishes, warm low latitude regions seem to not generate high metabolic rates and high substitution rates, as expected by the ESH, which is also the case for some terrestrial vertebrates (Jetz et al., 2012; Schluter & Pennell, 2017). There are also examples of crustaceans from cold and warm regions showing the same molecular substitution rate indicating no direct effect of temperature on molecular evolution (Held, 2001). Thus, our results suggest that latitude is not related to substitution rates in agreement with previous empirical results or that our approach was insufficient to detect general patterns. Average BL is probably more affected by mechanisms acting on the species level and does not escalate to upper levels originating a general response that corroborates the ESH. Animals from lower latitudes live closer to their upper thermal limit and have a low thermal tolerance (Vinagre et al., 2016). Because COI is a key gene in aerobic metabolism and species respond differently to changes in temperature, we may hypothesize that unique metabolic responses might emerge (e.g., Jost et al., 2012). Therefore, we may not detect a general pattern in COI substitution rates. Although we did not find support for higher substitution rates driving LGDG, GD is still higher on the lower latitudes.

Other explanations can be evoked to explain latitudinal patterns (reviewed in Mittelbach et al., 2007; Fraser et al., 2020). GD may peak in lower latitudes because this region has had more time to evolve, resulting in GD accumulation across species. Simultaneously, the tropics cover a significant part of Earth's surface, and larger areas can support larger population sizes, which tend to show higher GD. However, it is still unclear if population size influences molecular evolution rates (Dowle et al., 2013). Finally, biotic interactions can create opportunities for differentiation leading to higher GD (Vellend et al., 2005). Marine systems also may be under different forces compared to the terrestrial environment (Valentine & Jablonsky, 2015). The effect of currents seasonality, upwelling and monsoonal events in shallow systems and the drivers of deep-sea diversity represent challenges that remain unresolved

(Valentine & Jablonsky, 2015). There are still open questions about the factors influencing GD and the connections between intra- and interspecific GD, indicating that further studies are still needed as pointed out below (Fraser et al., 2020). It is feasible to assume that an interplay among forces is responsible for the patterns we have found. Here, we could not determine the processes behind it.

Even though we indicate LGDG, our work does not have a global sampling and a broad taxon sampling, mostly due to the lack of genetic information on crabs. Marine invertebrates, including crabs, are among the least studied organisms (Beheregaray, 2008). Indeed, a search on NCBI Taxonomy Database (Dec/2020) returned 576,578 nucleotides entries for "Brachyura" while "Mammalia" returned 81,086,209. Further studies should investigate GD spatial patterns in other marine groups (including those that show inverse LGD, like amphipods), different regions, expand the latitudinal range to temperate zones, use other molecular markers, and test the generality of our results. Until we do not have a comprehensive database for these neglected marine taxa, ocean-basin-scale, and taxon-based studies can help us to have a better appraisal of GD spatial patterns. Thus, our study represents a first step towards understanding GD spatial patterns and processes in marine crabs.

Idiosyncrasies explain intraspecific GD spatial patterns

Although a general LGDG has been found, our results on intraspecific data show species idiosyncratic patterns. Depending on the species, we found a different GD spatial distribution, and GD and BL were not always in accordance. Many species show no pattern, like *A. angustipes, U. cordatus, S. rectum, G. cruentata., A. petiverii, E. gonagra, C. danae, C. ornatus,* and *A. cribrarius,* which can be explained by the MH. Marine species can show large population sizes, and great distance dispersal can be achieved due to the presence of pelagic larva (Shanks, 2009). Many of the studied species show no genetic structure among their populations (Marochi et al., 2017; Buranelli & Mantelatto, 2017, 2019; Buranelli et al., 2019; Peres & Mantelatto, 2020). The combination of high gene flow and large population sizes can prevent any local trend from emerging, thus not following the proposals of other

GD spatial patterns hypotheses. Divergent haplotypes originated in low latitudes can be easily transported to other regions; therefore, we do not find a signal of ESH as haplotypes are not confined to warm temperature regions. Another case is that peripheral populations might be large and are continually incorporating genetic variation from other populations and not suffering from genetic drift effects, preventing the outcomes anticipated by the CMH. The species that follow the MH do not show any pattern regarding habitat or life-history traits. We have representatives of mangrove, rocky shore, and infralittoral species, so as species with high and low fecundity, indicating that the pattern we have found is probably generated by the high dispersal capability of these species. *Callinectes ornatus* and *L. leptodactyla* follow the MH, but their BL shows a different trend that we will discuss later.

Among the studied species, *A. pisonii* is the only one to show a clear pattern of higher GD on lower latitudes and the edges of its distribution. As this species shows a broad distribution encompassing both hemispheres, the LGDG and CMH would be confounded, as the higher GD is expected on lower latitudes, which is also the central region of the species distribution. However, BL also follows the same pattern as the GD. This result indicates that the ESH might be the mechanism behind the pattern, and the temperature influences substitution rates in this species, and *A. pisonii* shows a LGDG. The sudden higher GD and BL on the marginal ranges are somewhat counterintuitive, as theory suggests small population sizes at the edges of the distribution and consequently more affected by genetic drift leading to decreasing diversity (Eckert et al., 2008; Fraser et al., 2020). A possible scenario that can explain our result is population layering during the species range expansion (Schär et al., 2016). The edge of *A. pisonii* distribution might not show ideal conditions for establishing populations, which are probably composed of few individuals (Eckert et al., 2008). These populations might be continuously affected by harsh conditions and genetic drift. However, multiple cycles of expansion on the edges may result in different sets of genetic variation invading at each time, causing a sudden increase of GD and BL (Schär et al., 2016). Nearly-neutral theory predicts that slightly deleterious mutations are not eliminated in small populations due to relaxed purifying selection, which could also explain our results (Ohta, 1992). This scenario represents a complex interplay among many mechanisms acting on intraspecific GD and BL

spatial patterns, as we have temperature, migration, ecological constraints, and genetic drift acting upon *A. pisonii* (Buranelli & Mantelatto, 2019). Similarly, though *L. leptodactyla* does not show any GD spatial pattern, BL shows the same pattern as *A. pisonii* but less prominently. As we discussed, BL latitudinal trends might be caused by the temperature affecting substitution rates, and higher BL in marginal populations might be the effect of the accumulation of mutations in small populations. It is not clear why GD and BL do not follow the same pattern.

Minuca rapax shows higher GD on lower latitudes, but the GD decrease is different between hemispheres, being less sharp on the North Hemisphere. As in *A. pisonii*, this pattern could be the result of LGDG or CMH. In this case, we do not find support for higher substitution rates resulting in LGDG, and BL depicts a linear trend along latitude. The same linear trend for BL is also found in *N. granulata*. However, *N. granulata* is confided in the southern hemisphere. We see a bell-shape distribution when we look at the GD spatial pattern in *N. granulata*, showing a peak in the middle of the species distribution, which was already demonstrated in the literature but not discussed (Ituarte et al., 2012). The pattern found for *N. granulata* matches precisely what the CMH predicts. As both *M. rapax* and *N. granulata* show the same BL trend, we believe *M. rapax* trend is also the result of the CMH instead of LGDG. Thus, both species show higher GD on the center of their distribution, but they probably depict different GD distribution shapes due to not being co-distributed (*M. rapax*: the Caribbean and Tropical South Atlantic; *N. granulata*: Temperate South Atlantic). We interpret their linear BL trend as continuous population expansion and retractions towards South latitudes resulting in population layering. This process seems to explain the pattern found in five butterfly species from temperate zones, in which it results from population dynamics during glacial and interglacial cycles (trans-glacial latitudinal layering hypothesis - Schär et al., 2016). In our case, the crab species were not affected by glaciations but probably affected by sea-level fall (Toms et al., 2014). We believe our results show a series of expansions and retractions (Schär et al., 2016), mainly because coastal species can be directly affected by sea-level changes (Marko et al., 2010).

Leptuca thayeri is the only one to show a dip in GD on lower latitudes and peaks on intermediate latitudes depicting a bimodal distribution. This type of distribution also characterizes the marine species richness trend (Chaudhary et al., 2016; Chaudhary et al., 2017), also found in other crustaceans (Rivadeneira & Poore, 2020). In all cases, it is not clear what factors shape the bimodal distribution, but probably it is caused by a combination of factors. Moving to GD, it is also not clear why we found an intraspecific GD bimodal distribution. Interestingly, "Uca" crabs show bimodal species richness distribution (Levinton & Mackie, 2013), so we may hypothesize that the same factors also affect *L. thayeri* GD, but not the other "Uca" crabs we examined. For some marine species, connectivity at intermediate latitudes is lower than at other latitudes (Álvarez-Noriega et al., 2020). Hence, we may find higher population differentiation and, ultimately, higher GD in these populations, an assertion that is pending testing. *Leptuca thayeri* shows a strong genetic structure showing two distinct genetic units (Buranelli & Mantelatto, 2019). An alternative explanation for our results is that these two groups behave as separate units and GD spatial patterns follow the CMH. Under this scenario, we would expect to find higher GD on the central range of each unit distribution, which agrees with our results. On the other hand, BL does not follow the same trend as the GD and shows higher BL at the Caribbean side of the distributional range of the species. The same BL pattern was also found in *C. ornatus*, a species that also shows two genetic clusters (Peres & Mantelatto, 2020). Our results may represent the phylogeographical diversification scenario that took place in both species. Long BL are associated with recent haplotypes (Posada & Crandall, 2001), so we can hypothesize that a dispersed event occurred from Southern Western Atlantic populations to the Caribbean, and new populations were founded in the past, probably after the Pleistocene (Maggs et al., 2008). Although Tropical regions did not show glacial sheets, the sea level fell significantly, affecting coastal species (Toms et al., 2014), and recolonization events may occur as in temperate species (Maggs et al., 2008). The sudden increase in BL might result from sampling bias, as some intermediate populations were not sampled. We believe that if present, we would find a linear trend as shown for *M. rapax* and *N. granulata* but in the opposite direction (South-North direction).

It is important to notice that among the species showing GD spatial patterns, they also show some kind of geographic genetic structure (Ituarte et al., 2012; Laurenzano et al., 2016; Buranelli & Mantelatto, 2019; Peres & Mantelatto, 2020). Some species that follow the MH show different genetic clusters but are not related to geographic structure (e.g., Buranelli & Mantelatto, 2017). This can indicate that GD spatial patterns may only emerge when there are constraints to the larval dispersal. We may think the MH as a null hypothesis that assumes high gene flow among populations, but when any factor prevents the lack of geographic genetic structure, some GD spatial patterns may emerge.

Finally, our findings are timely as we face a biodiversity crisis (Barnosky et al., 2011; Lewis & Maslin, 2015). The United Nations Convention on Biological Diversity (CBD; www.cbd.int) considers the GD as one of level of diversity that we should strive to preserve. However, GD has been consistently neglected during the implementation of strategies targeting management and conservation (Laikre, 2010). Although commonly discussed at the intraspecific level, GD directly affects communities and ecosystems by influencing, for instance, productivity, decomposition, pollination, and consumer-resource dynamics (Hughes et al., 2008). Here, we demonstrate that crabs show average higher GD on lower latitudes, indicating that preserving these regions might be more effective than preserving others, as suggested for terrestrial vertebrates (Adams & Hadly, 2013). The marine environment still faces challenges as we lack information on GD, species richness patterns, and population/community dynamics, which are crucial for designing and evaluating marine protected areas (Fox et al., 2014; Sandström et al., 2019). Thus, it is imperative to increase our understanding of GD's patterns and processes in the marine realm.

Concluding remarks

One of the most striking patterns is the latitudinal distribution of diversity, and yet the processes causing such diversity remain unresolved. Our results show interspecific LGDG in western Atlantic crabs, but species GD spatial patterns seem to be species-specific. The MH, LGDG, CMH offers interesting explanations for the intraspecific level, but it is not clear how their combination resulted in interspecific LGDG. As we start to get more empirical data on GD spatial patterns across different species and environments, the more it seems these patterns are governed by an interplay among forces and might be context dependent. Common responses might be resulted from different mechanisms and deserve caution when interpreting results. It is important to consider that we restrict our analyses to latitude. Still, longitude and depth might also play a role in determining GD spatial patterns, and further studies should also consider these variables. In conclusion, our study offers empirical results of GD spatial trends at different organization levels. Combining species data allowed us to reveal trends that are not in accordance with single-species data, neither apparent from intraspecific analyses highlighting the complexity of GD geographic distribution.

CHAPTER 4 (Manuscript *in prep*.)

The determinants of the genetic diversity in crabs

Pedro A. Peres & Fernando L. Mantelatto

Introduction

Understanding what drives species genetic diversity (GD) is still one of the open questions in ecology and evolution (Leffler et al., 2012; Ellegren & Galtier, 2016). Paradoxically, we do not entirely understand how and why GD varies across species and, at the same time, it is recognized by the Convention on Biological Diversity (CBD, https://www.cbd.int/) as one of the levels of biodiversity to be conserved and sustainably used. GD is the trait responsible for species' adaptive potential when facing environmental changes (Eizaguirre & Baltazar-Soares, 2014) and for preventing endogamic depression due to the accumulation of deleterious mutations (O'Grady et al., 2006). GD also has effects above the population or species level, correlating with community composition and ecosystem functioning (e.g., productivity, decomposition) in terrestrial and marine environments (Hughes et al., 2008; Whitlock, 2014; Jormalainen et al., 2017). Thus, it is necessary to gain a deeper understanding of the determinants of GD.

The neutral theory of molecular evolution predicts that GD is proportional to the effective population size (N_e) at neutral sites because of the mutation/drift equilibrium. Hence, the bigger the N_e, the bigger the GD. The N_e may fluctuate due to environmental disturbances (i.e., climatic changes, biotic interactions, anthropic effects) increasing or decreasing populations and subsequently altering GD (Hewitt, 2004; Wenger et al., 2011; Banks et al., 2013; Lewis & Maslin, 2015; Gurgel et al., 2020). A significant impediment to estimate current N_e, and consequently its effect on GD, is that we would need the number of males and females, their genetic contribution to the next generation, besides meeting many assumptions (Wang et al., 2016). Thus, N_e approximations based on genetic parameters are often used to estimate the historical N_e, representing the harmonic mean of N_e over time, providing a proxy for N_e to test the relationship between this parameter and GD (Ellegren & Galtier, 2016; Wang et al., 2016).

This relationship has been tested empirically with different markers throughout the years (Soule, 1976; Frankham et al.,1996; Montgomery et al., 2000; Romiguier et al., 2014; Mackintosh et al., 2019). The population size-GD relationship usually indicates that groups known to have larger population sizes show higher nucleotide diversity (e.g., insects > mammals) (Frankham et al., 1996; Leffler et al., 2012). However, the relationship between GD of mitochondrial DNA (mtDNA) markers and population size remains ambiguous. There is support for the evidence of mtDNA GD and population size being proportional (Mulligan et al., 2006; Nabholz et al., 2008a; Piganeau & Eyre-Walker, 2009), so as for no relationship at all (Bazin et al., 2006), leaving this question demanding further investigation.

Life-history traits have also been shown as determinants of the GD across different taxa (Romiguier et al., 2014; Kort et al., 2021). Results from a comparison of 31 families from different phyla have shown that the combination of adult size, body mass, maximum longevity, adult dispersion ability, fecundity, and propagule size explained more than 70% of the GD variation (Romiguier et al., 2014). Interestingly, propagule size was the primary factor influencing GD. The authors argue that the contrast between r-K strategies (Pianka, 1970) allows K-strategy species to resist disturbances, holding stable or even small populations, while r-strategy species can hold larger populations but with higher risks of going through demographic changes when facing the same disturbances (Romiguier et al., 2014). A more extensive comparison, but restricted to mammals, birds, reptiles, amphibians, and mollusks, also found body size, longevity, and fecundity as important life-history traits depending on the taxa (Kort et al., 2021). In both cases, these traits are potentially related to N_e, offering a more accurate way to investigate the intensity of demographic changes over time that resulted in the GD presently found.

Studies focusing on one particular group instead of comparing highly divergent taxa could help to elucidate the factors influencing the GD (Leffler et al., 2012). The mutation rate for both nuclear DNA (nuDNA) and mtDNA is highly variable across animals, potentially affecting the N_e estimates during comparison among animal groups (Allio et al., 2017). Thus, investigating the determinants of GD in species sharing an evolutionary history but still showing variable traits across species could alleviate such problems. Also, current knowledge on GD is predominantly based upon vertebrate taxa, and we

might be missing new trends due to the lack of investigation on neglected taxa. For instance, bony fishes show a negative relationship between GD and maximum size, egg diameter, and length at maturity as expected (Mitton & Lewis Jr, 1989), but some butterfly families, though showing a negative correlation between GD and size, show no relationship between GD and egg size, larval host plat and current abundance (Mackintosh et al., 2019).

Habitat has also been shown as an important factor for GD. Upland Amazonian bird species show higher GD than floodplain species (Harvey et al., 2017), terrestrial birds show higher GD than aquatic birds (Eo et al., 2011), marine fishes show higher GD than freshwater species (DeWoody & Avise, 2000; Martinez et al., 2018), and shallow decapod species show higher GD than deep-sea species (García-Merchán et al., 2012). Similar habitats may undergo the same geological and abiotic changes leading to similar demographic responses influencing the GD of the habitat-associated fauna in terrestrial and marine environments (Marko et al. 2010; Gehara et al., 2017). Also, some habitat types are more connected through species dispersal resulting in patterns like canopy bird species being less genetically differentiated than understory species (Burney et al., 2009) and less genetic differentiation explained by depth in marine animals (Etter et al., 2005; García-Merchán et al., 2012; Selkoe et al., 2014). In many marine species, dispersal occurs through a planktonic larva that remains in the water column and may be transported by currents (Shanks, 2009). For these species, the dispersal ability can be related to the number of larval stages and the pelagic larval duration (PLD), promoting connectivity among populations (Faurby & Barber, 2012). Population connectivity might hamper the erosion of the GD caused by genetic drift and population size decrease by inputting new individuals into these populations. Yet, there might be differences between the potential and realized dispersal (Weersing & Toonen, 2009), and we still lack the use of the number of larval stages and PLD as predictors of the GD (Kort et al., 2021).

Considering the open questions on the determinants of GD, the benefits of investigating related groups, and the need to explore GD patterns in neglected taxa, crabs (Brachyura) emerge as a model taxa. Brachyura are one of the most diversified invertebrates and one of the most studied crustaceans

(Ng et al., 2008; Davie et al., 2015; Wolfe et al., 2021). Crabs are found from abyssal zones to terrestrial environments occupying most habitats, showing a vast life-history traits variation and dispersal potential (Hines, 1982; Hines, 1986; Ng et al., 2008; Anger et al., 2015; Davie et al., 2015). There are indications that species density, fecundity, and demographic changes explain the GD for seven mangrove crab species from the Western Indian Ocean (Fratini et al., 2016). A broad taxonomic sampling investigating species from different habitats and showing diverse life-history traits could unveil if this is a general trend in crabs and contribute to understanding what affects GD in an invertebrate group (see also Mackintosh et al., 2019).

Here, we investigated the determinants of the GD in crabs (Brachyura) by compiling the most comprehensive genetic dataset (cytochrome c oxidase subunit I – COI) to date. We focused on the mtDNA due to its popularity and confirmed applicability in barcode, e-DNA, metabarcoding, phylogenetic, and phylogeographic studies (Avise et al., 2000; Hebert et al., 2003; Lefébure et al. 2006; Matzen da Silva et al. 2011; Timm & Bracken-Grissom, 2015; Mantelatto et al. 2020; Collins et al., 2018). Also, the use of the COI gene is motivated due to its thousands of publicly available sequences (Porter & Hajibabaei, 2018) that are being used to investigate the determinants of GD in other taxa (Miraldo et al. 2016; Manel et al. 2020; Theodoridis et al. 2020). We tested the effect of different life-history and demographic variables (fecundity, size, propagule size, number of larval stages and larval development time, historical effective population size, maximum longevity, habitat) that potentially influence GD. The variables were chosen to better represent the main aspects of species life history, like population size, their dispersion ability, and ecological strategies, although we are aware that some of them are difficult to disentangle from one another. We hypothesize that (1) traits related to population size show either a negative correlation (body size, maximum longevity) or a proportional association (historical N_e) to the GD based on the theoretical relationship between population size and GD; (2) traits related to dispersion ability (number of larval stages and larval development time) show a positive correlation to GD by spreading alleles across populations and potentially hampering the effect of genetic drift; (3) traits related to ecological strategies (fecundity and propagule size) show a negative correlation

to GD due to the impacts of r/K strategies; (4) GD varies across habitat because connectivity might be dependent of habitat, and because similar habitats might have undergone similar environmental changes

Materials and Methods

Species sequence collection

Brachyura is formed by the sister clades Podotremata and Eubrachyura (Ng et al., 2008; Guinot et al., 2013; Wolfe et al., 2019). We used the Taxonomy Browser tool in NCBI (https://www.ncbi.nlm.nih.gov/taxonomy) to first search all Brachyura genetic sequences publicly available. However, Podotremata crabs were highly underrepresented and did not meet our criteria for retrieving sequences (see below). Hence, they were not included in our dataset. We detected inconsistencies in how COI sequences are named in the database, leading to different sequence sets retrieved depending on the name used for searching. Thus, we searched the terms “cytochrome c oxidase subunit 1”, “cytochrome oxidase subunit 1”, “cytochrome oxidase 1”, “COI”, and “COX1” within Eubrachyura to ensure a total inclusion of COI data. We also noticed that many authors submitted only unique haplotypes, even though they were originated from populational studies, which could potentially influence our GD estimates. Initially, we retrieved all species showing more than two sequences because they could represent unique haplotypes from a larger dataset and recovered 210 crab species. Then, we reconstructed the species' haplotype frequencies that had available only unique haplotypes by consulting the reference article or contacting the authors. We could not reconstruct haplotype frequencies for the species with sequences deposited with no reference article and were thus excluded from downstream analyses. This step allowed us to have a better appraisal of the intraspecific genetic variation.

We decided to keep in our dataset only the species with more than 15 sequences in the database or that had more than 15 sequences after haplotypes frequency reconstruction in the cases mentioned before. This threshold was chosen based on simulation and empirical results demonstrating that 15 sequences could ensure us to have a comprehensive picture of the GD within species (Goodall-Copestake

et al., 2012; Luo et al., 2015; Phillips et al., 2019). We followed the authors' interpretation who generated the sequences when they found cryptic species through molecular data but not formally described them. These species were maintained as separate species in our dataset (i.e., if the authors addressed more than one species under the same species name, downstream analyses used the sequences for each cryptic species). For each species, sequences were aligned using MAFFT v.7 (Katoh & Standley, 2013) and visually inspected in Geneious Prime 2020.2.4 (https://www.geneious.com). Within each species alignment, sequences shorter than 375 base pairs (bp) or showing incongruences with the rest of the alignment were excluded.

Response variable - Genetic diversity (GD)

We estimated the GD of each species using the nucleotide diversity (π) calculated in DnaSP v.6 (Rozas et al., 2017). Nucleotide diversity is one of the measures of the GD and represents the average number of nucleotide differences per site between two sampled DNA sequences (Nei & Li, 1979). After this step, we excluded species that showed $\pi > 0.02$ due to the potential presence of cryptic species in the dataset, which inflated π estimates beyond the values found for the majority of our species and are also considered above intraspecific π for other taxa (Goodall-Copestake et al., 2012).

Predictor variables - Life-history and demographic traits

We investigated the influence of fecundity, egg diameter (propagule size), size (maximum carapace width - CW), number of larval stages, larval development until crab 1 phase in days (pelagic larval duration - PLD), maximum longevity in years, habitat, and historical effective population size (N_e) on the GD. Using Google Scholar©, we searched for articles using the species name combined with the terms “fecundity” OR “population” OR “larva” to obtain data for fecundity, egg size diameter, maximum CW, larval development, and habitat. All data were retrieved from scientific articles in

English. We did not include grey literature, dissertations, thesis, non-scientifical articles, or scientific articles in other languages. When multiple papers containing a targeted variable were available, the trait's value used resulted from the average value among all articles (e.g., average among the fecundity for different localities; an average of larval development days under different temperature or salinity treatments). This was done to get a unique value to represent the species trait while also considering its variation. If necessary, egg volume (v) was transformed to egg diameter using the equation $v = (\pi \times \text{diameter}^3)/6$ (Hines, 1982; Terossi et al. 2010; Peres et al. 2018). Maximum longevity was in most case retrieved from Vogt (2019). The number of larval stages was obtained from the larval development articles and, when necessary, extrapolated to species with no available information from the same genus. Habitat was classified in the categories deep-sea, symbiotic, infralittoral, intertidal, mangrove, estuarine, and terrestrial. We also estimated the Watterson's θ (Watterson, 1975) in DnaSP v.6 to calculate the historical N_e using the equation $\theta = 2\ N_e\mu$ for mitochondrial DNA, where μ is the mutation rate per sequence per generation. We considered $\mu = 0.01 \times 10^{-6}$ because this is the standard mutation rate for mtDNA used for decapods (Knolwton et al. 1993).

Statistical analyses

Because our dataset contained missing values, we followed two different strategies. First, linear regression models were performed between π and each continuous predictor variable (fecundity, egg diameter, maximum CW, number of larval stages, PLD, maximum longevity, historical N_e) and 1-factor-ANOVA between π and habitat (categorical variable) followed by posthoc Tukey's test to investigate the relationship between them using data from the highest number of species. The strategies allowed us to maximize the amount of data for each variable.

In the second strategy, we performed multiple linear regression models followed by model selection using the information-theoretic approach (Burnham & Anderson, 2002). A multimodel information-theoretic approach is preferable over stepwise and backward model selection because it

considers all possible variable combinations, and the importance of variables can be explored while considering model uncertainty (Johnson & Omland, 2004; Mundri & Nunn, 2009). Multiple linear regressions were performed using 13 sets of species and variables (Table 1). We adopted this strategy because our full dataset contained 27 species, as these were the only ones who had all seven predictor variables available. Thus, we decided to exclude some variables and test other 12 sets (13 in total) to maximize the number of species included in each tested model to have more confidence in our results. Because maximum longevity is our least sampled variable, we restricted it to just one model combined with historical N_e (the most sampled variable). Also, freshwater species were not used in most analyses because we just had data for habitat and historical N_e (all species) and maximum longevity for three species, but no information on any other predictor variables. Thus, all of our discussion is mainly based on marine species. For each set, we used the *dredge* function available in the package MuMIn (Barton, 2020). This function fits models for subsets of the global model, performing an automated model selection for all possible combinations of predictor variables. These models are then ranked based on AICc, and models with $\Delta AICc < 2$ were considered the best ones explaining the relationship among the response and predictor variables present in the top subsets (Burnham & Anderson, 2002). When more than 1 model had $\Delta AICc < 2$, we used the function *model.avg* to do model averaging and estimate the importance of each predictor variable based on Akaike Weights (w). A variable showing $w = 1$ indicates that the variable was present in all candidate models (Burnham & Anderson, 2002). Standardized partial slope coefficients are reported with the 95% confidence intervals (CI), and significant effects were accepted when CI did not include zero. Before running the analysis, the predictor variables were centered, and we checked for collinearity among our variables using Spearman's rank-order correlations (ρ) and the variance inflation factor (VIF and 1/VIF). In all sets, fecundity and maximum CW showed signs of collinearity ($\rho > 0.7$; $VIF < 3$; $1/VIF > 0.2$). When both variables were present, we first did a PCA to create a reduced set of orthogonal variables and used the PC1 in the multiple regression and model selection.

Before both analyses, the normality, homoscedasticity, and independence of residuals were visually inspected on residual-plots (Boldina & Beninger, 2016). Outliers were removed, or the data were log-transformed when necessary. All analyses were run in R software (version 4.0.2).

Table 1. Summary of all datasets tested. Each dataset contained a different set of variables. Max CW: maximum carapace width; PLD: pelagic larval duration; Ne: effective population size.

Dataset	Variables	Number of species
1	Fecundity + Max CW + Egg diameter + Number of larval stages + PLD + Historical Ne	27
2	Fecundity + Max CW + Egg diameter + Number of larval stages + Historical Ne	47
3	Max CW + Egg diameter + Number of larval stages + Historical Ne	46
4	Fecundity + Egg diameter + Number of larval stages + Historical Ne	47
5	Fecundity + Max CW + Egg diameter + Historical Ne	43
6	Egg diameter + PLD + Historical Ne	30
7	Egg diameter + Number of larval stages + Historical Ne	51
8	Fecundity + Max CW + Historical Ne	60
9	Egg diameter + Historical Ne	52
10	PLD + Historical Ne	53
11	Max CW + Historical Ne	91
12	Number of larval stages + Historical Ne	128
13	Historical Ne + Max longevity	28

Results

After the filtering steps, we ended up with a dataset of 150 species and 16992 COI sequences. The dataset encompasses 85 genera, 33 families, and 18 superfamilies. We retrieved 64 species data for fecundity, 91 for maximum CW, 52 for egg diameter, 128 for the number of larval stages, 53 for PLD, 31 for maximum longevity, 150 for historical N_e, 150 for habitat. Across all crabs used (150 species), the largest π (0.01868) is approximately 25x the smallest π (0.00074). The π variance we found is not a result of the number of sequences used per species, as shown by the lack of relationship among these variables (linear regression: $F = 1.452$, $p = 0.2302$, $R^2 = 0.00303$, data not shown). Regarding linear regressions, three variables showed significant relationships with π: egg diameter, maximum longevity,

and historical N_e (Table 2, Figure 1). Species showing smaller eggs tend to have higher genetic diversity, while long-living species and species showing higher long-term population sizes have higher genetic diversity. Fecundity, maximum CW, number of larval stages, and PLD do not affect genetic diversity (Figure 1). Regarding habitat, we found statistical differences among them (ANOVA: $F = 2.797$, $df = 7$, $p = 0.0093$), but the posthoc Tukey's test could not detect the differences (Figure 2).

Table 2. Linear regression results for each variable tested. Significant p-values are indicated in bold

Variable	F	p-value	R2
Fecundity	0.321	0.5737	0.005397
Maximum Carapace Width	1.776	0.1862	0.002
Egg diameter	4.78	**0.03416**	0.07749
Number of larval stages	2.211	0.1399	0.01917
PLD	3.006	0.08998	0.06394
Historical effective population size	116.8	**<0.0001**	0.4565
Maximum longevity	8.715	**0.008184**	0.2784

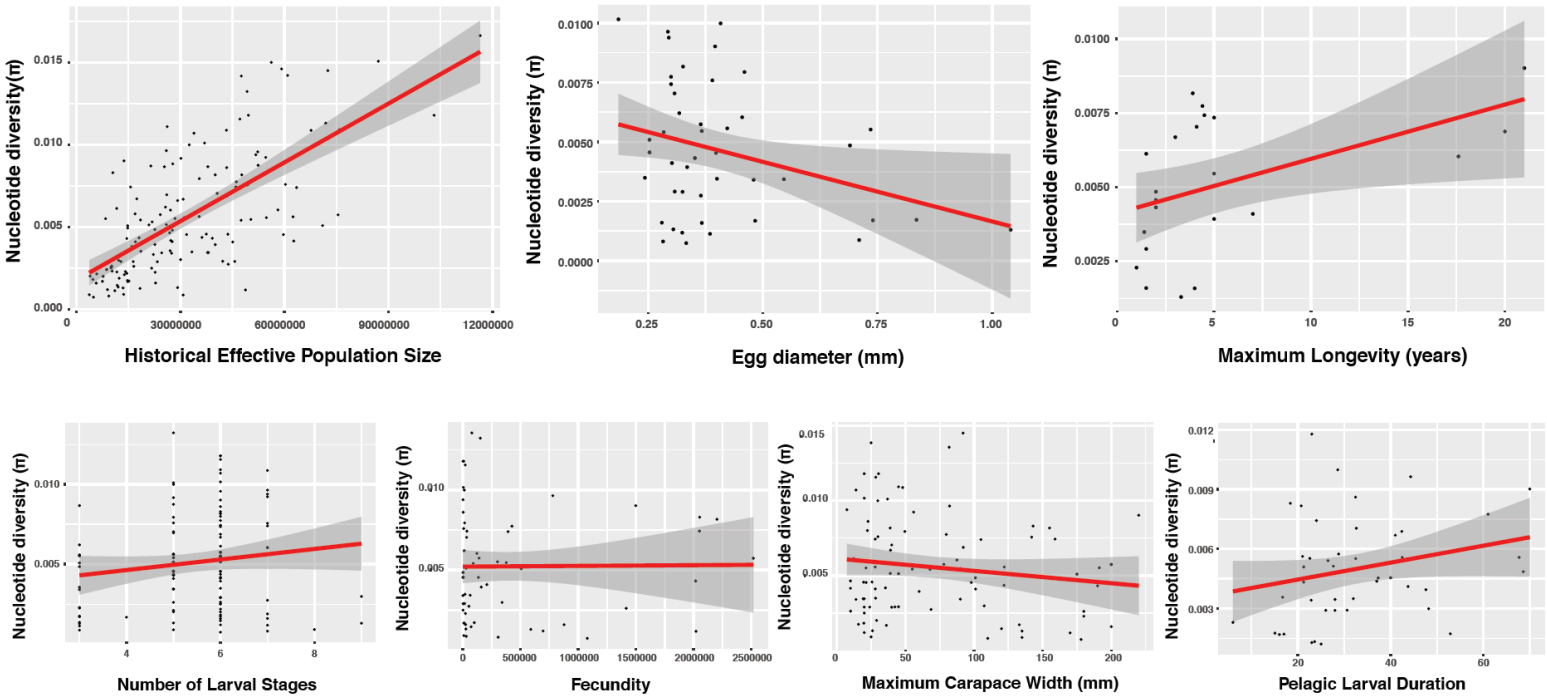


Figure 1. Linear regression graphs of each variable. Graphs on the upper part depict significant associations, while graphs on the bottom part are not significant. The gray area around the red trend-line represents the standard error.

Different sets of variables were included in the best models after model selection in our multiple linear regression analyses, with variables showing distinct relative importance depending on the dataset tested (Table 3). Historical N_e was found as a determinant variable in 11 of the datasets, PC1 (datasets combining maximum CW and fecundity) and PLD in 3, egg diameter in 2, fecundity and maximum CW in 1 (datasets testing the variables separately), and maximum longevity in 1. However, significant effects were just found for Historical N_e, PC1, PLD, and egg diameter depending on the dataset tested. The multiple regression analyses indicate small, low fecund, showing smaller eggs, long PLD, and larger effective population size species tending to have higher GD. Overall, historical N_e had the higher relative importance (w = 1) when present in the top models in the datasets tested, explaining the relationship with genetic diversity.

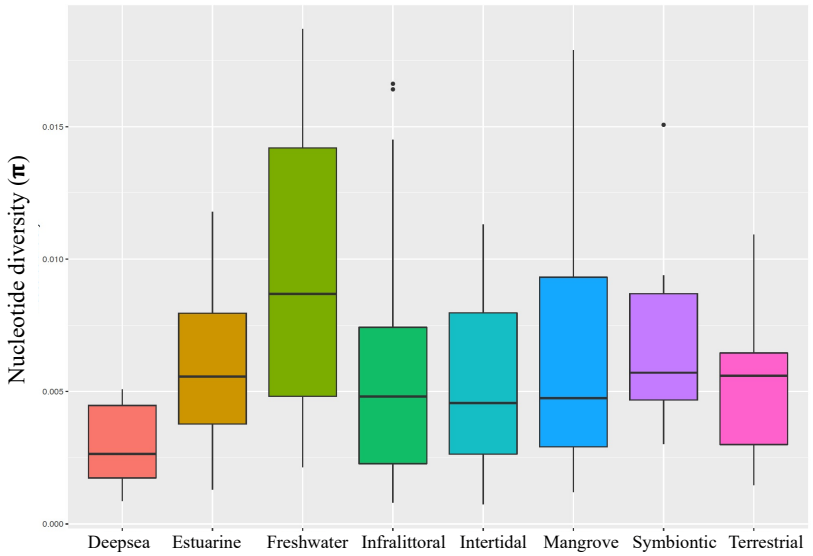


Figure 2. Comparison among genetic diversity across different habitats.

Table 3. Top selected model or model-averaged standardized regression coefficients (β), 95% confidence intervals (CIs), and weight (w) for each dataset tested. Significant variants based on Cis are indicated in bold. n: number of species contained in the model; Max CW: maximum carapace width; PLD: pelagic larval duration; N_e: effective population size.

Dataset	Number of selected models	Predictor variable	β	Lower CI	Upper CI	Weight (w)
1 Fecundity + Max CW + Egg diameter + Number of larval stages + PLD + Historical N_e n = 27	1	Egg diameter	-0.50131	**-1.46E-02**	**-1.96E-03**	
		PC1	-0.51883	**-1.53E-03**	**-3.06E-04**	
		PLD	0.50718	**1.75E-05**	**1.10E-04**	
2 Fecundity + Max CW + Egg diameter + Number of larval stages + Historical N_e n = 47	2	Historical Ne	6.58E-01	**5.52E-11**	**1.19E-10**	1
		PC1	-7.87E-02	-9.13E-04	4.65E-04	0.27
3 Max CW + Egg diameter + Number of larval stages + Historical N_e n = 46	1	Historical Ne	6.58E-01	**5.63E-11**	**1.17E-10**	
4 Fecundity + Egg diameter + Number of larval stages + Historical N_e n = 47	2	Historical Ne	6.45E-01	**5.72E-11**	**1.22E-10**	1
		Fecundity	9.26E-02	-7.23E-10	1.66E-09	0.3
5 Fecundity + Max CW + Egg diameter + Historical N_e n = 43	2	Historical Ne	6.58E-01	**5.52E-11**	**1.19E-10**	1
		PC1	-7.87E-02	-9.13E-04	4.65E-04	0.27
6 Egg diameter + PLD + Historical N_e n = 30	4	Egg diameter	-3.50E-01	-0.011148047	3.77E-04	0.64
		PLD	3.31E-01	-6.1793E-06	9.92E-05	0.6
7 Egg diameter + Number of larval stages + Historical N_e n = 51	1	Historical Ne	7.04E-01	**6.07E-11**	**1.14E-10**	
8 Fecundity + Max CW + Historical N_e n = 60	1	Historical Ne	6.07E-01	**5.67E-11**	**1.17E-10**	
9 Egg diameter + Historical Ne n =52	1	Historical Ne	6.56E-01	**6.08E-11**	**1.21E-10**	
10 PLD + Historical N_e n = 53	2	Historical Ne	3.82E-01	**1.77E-11**	**1.00E-10**	1.00E+00
		PLD	2.45E-01	-2.87E-06	8.00E-05	0.65
11		Historical Ne	0.550594	**5.76E-11**	**1.12E-10**	1.00E+00

Max CW + Historical N_e n = 91	2	Max CW	-6.37E-02	-1.53E-05	7.18E-06	0.3
12 Number of larval stages + Historical N_e n = 128	1	Historical Ne	6.53E-01	**7.85E-11**	**1.21E-10**	
13 Historical Ne + Max longevity n = 28	1	Historical Ne	8.89E-01	**1.02E-10**	**1.62E-10**	1
		Max longevity	0.10516	-1.24E-03	3.75E-03	0.3

Discussion

Our results obtained from eight predictor variables interacting with 150 species support the hypothesis that species effective population size and ecological traits can predict patterns of genetic diversity (GD) for crab species. Our most consistent evidence shows the historical effective population size (N_e) as the most determinant factor predicting GD. N_e stands out as a high-explanatory and significant variable in both linear and multiple linear regressions, showing proportional effects on the GD. Other variables show effects on the GD depending on the analyses, like egg-diameter (propagule size), maximum longevity, fecundity, size, and pelagic larval duration (PLD). GD also seems to be different across habitats, although we did not have precise results. As far as we know this is the first work that focus on combining life history traits to better understand the determinants of GD providing results from a neglected yet diverse group of animals - the crabs (Eubrachyura).

Traits related to population size (historical N_e, body size, maximum longevity)

Direct estimations of the N_e might be complex (Wang et al., 2016), but we can use other variables as proxies to explore the effect of N_e on the GD, such as body size, maximum longevity, and historical N_e. Species body size is considered to be negatively associated with species abundance, census population size, and mutation rates, consequently negatively correlated to GD (Wooten & Smith, 1985; Martin & Palumbi, 1993; Eo et al., 2011). Therefore, smaller species would show higher GD than larger

species because they are more abundant, show larger census population size, more variability is maintained by different microhabitat selective pressures, and show higher mutation rates (Selander & Kaufman, 1973; Martin & Palumbi, 1993). In our analyses, body size (represented by maximum carapace width - CW) was found as an important variable in dataset 1 (n = 27, predictor variables: fecundity, maximum CW, egg diameter, PLD, number of larval stages, historical N_e) and dataset 11 (n = 91, predictor variables: maximum CW, historical N_e). Although selected in the best explanatory models, it was significant in just one case. Similarly, the relationship found in linear regression between maximum CW and GD was also not significant.

Body sizes show a negative relationship to GD in many mammals, birds, and fish species (Wooten & Smith, 1985; Mitton & Lewis Jr, 1989; Eo et al., 2011; Kort et al., 2021), but some studies also do not support this affirmation (Doyle et al., 2015; Azizan & Paradis, 2021; but see Kort et al., 2021 which includes mollusks in their dataset). Body size was significant just in the dataset showing the least number of species (n = 27), raising doubts if this is, in fact, a real trend. Our results might represent a taxon-specific lack of body size-GD because there is evidence that mtDNA and nuDNA are proportional (Mulligan et al., 2006; Nabholz et al., 2008a; Piganeau & Eyre-Walker, 2009), refuting the idea of mtDNA is not capturing the population size variation (Bazin et al., 2006). GD calculated from nuDNA is not correlated to body size in Felidae (Azizan-Paradis, 2021) and other mammals (James & Eyre-Walker, 2020), reinforcing the idea of a possible taxon-specific lack of body size-GD correlation. Also, many of the larger species are from the infralittoral, an enormous habitat compared to delimited ones like mangroves and rocky shores, allowing the existence of large populations even for large body animals. Therefore, large animals can sustain large populations in specific habitats leading to a pattern that does not follow initial predictions (Wooten & Smith, 1985; Martin & Palumbi, 1993; Eo et al., 2011).

Longevity is expected to be negatively correlated to GD as it is often associated with large species, small-size populations, and slower mutation rates (Martin & Palumbi, 1993; Nabholz et al., 2008b). Multiple taxa comparisons indeed find this association (Romiguier et al., 2014; Chen et al.,

2017; Kort et al., 2021); however, longevity often shows no relationship to GD within a taxon (Mitton & Lewis Jr, 1989; Nabholz et al., 2008a; Mackintosh et al., 2019). We found an opposite effect and show a proportional and significant association between longevity and GD in the linear regression analysis, yet no significance in multiple models (dataset 13). We choose to interpret our results as no association between maximum longevity and GD. The significant simple model represents our least sampled trait (n = 31), which we believe has biased this result. A lack of association between longevity and GD aligns with the body size results. Although a general trend exists across many groups, crabs might show a taxon-specific pattern that might also be explained by larger habitats sustaining larger populations of larger animals. In the future, when more longevity data is available, this association should be explored again in a more comprehensive view.

Finally, historical N_e, the last trait related to population size, was the most important predictor variable explaining GD. In the linear model, the variation in historical N_e explains 45% of the GD variation. In multiple linear models, historical N_e was present in the top models in 11 out of 13 datasets tested and significant in all cases. By assessing historical N_e, we estimate the mean N_e variation throughout generations as if N_e remained constant over time, an unlikely assumption, so recent or past demographic changes are not being accounted for (Waples, 2005; Wang et al., 2016). Although we are not estimating contemporary N_e, historical N_e might be more important to capture the accumulation of GD over time despite the demographic changes that had occurred (Ellegren & Gaultier, 2016). Our results show that species that maintained larger populations over time are more genetically diverse, as expected by the neutral theory of evolution. Larger populations are less affected by genetic drift and consequently do not lose GD at faster rates like smaller populations, sustaining higher GD. Larger populations also have more chances of new mutations arising, contributing to higher GD. The use of mtDNA to estimate N_e is controversial because it is prone to genetic hitchhiking due to its lack of recombination, which affects the assumption of neutrality (Ballard & Kreitman, 1995; Galtier et al., 2009). As we mentioned before, this argument is questionable, and many studies indeed show a relationship between neutral nuDNA and mtDNA (Mulligan et al., 2006; Nabholz et al., 2008a Piganeau

& Eyre-Walker, 2009). Even if selection acts on mtDNA, it might not be the dominant evolutionary force affecting variation, and we can assume the mtDNA as nearly neutral (Ohta, 1992; Figuet et al., 2016), and the mtDNA-Ne association remains reliable. A shortfall of our estimates is that we assumed a constant substitution rate (μ), which might not be accurate. Even though this parameter can vary across different lineages (Nabholz et al., 2008b; Silva et al., 2011), we believe it is unlikely that the parameter variation could disrupt our result making historical N_e and GD not showing any association, especially considering the effect size we found, and we would only see, if any, changes in the effect size and not in the direction or significance of the relationship. That is, we would still find historical N_e showing a highly explanatory power on GD. Therefore, we are confident our analyses captured the effect of N_e over GD being the most crucial determinant factor.

Our results have important implications on species conservation because we show that low N_e is usually associated with low GD. GD is related to the potential of a species to cope with environmental changes and also reduces the effects of deleterious mutations (Eizaguirre & Baltazar-Soares, 2014). Examining the top 20 lowest N_e in our dataset, we did not find a genus, family, superfamily, or habitat trend, indicating that other factors might be driving the low GD. Different forces can act upon species depending on their location or traits and demand further investigation. Local historical habitat changes, biotic interactions, anthropogenic impacts, or even reproductive patterns can drive changes in GD (Hewitt, 2000; Hughes et al., 2008; Marko et al., 2010; Wang et al., 2016; Schlaepfer et al., 2018). For instance, species such as *Cancer pagurus* can be targeted in future studies to explore the relationship between low N_e andreproductive behavior. Although not in the top 20 lowest Ne (but in the top 30 species), *C. pagurus* females' mate with multiple males during the reproductive season, but the brood is all from a single partner (McKeown & Shaw, 2008). Such reproductive behavior can have impacts on individuals' genetic contribution to the next generation and, consequently, in the N_e (Nunney, 1993; Wang et al., 2016)

Traits related to ecological strategies (fecundity, egg diameter)

The r/K strategies have been evoked to explain GD variation across many species (Romiguier et al., 2014; Chen et al., 2017). They represent different life-history alternatives, being r-strategists the species showing high fecundity, short life-span, early reproduction, low parental investment in the offspring, small body size; in contrast, K-strategists show low fecundity, long life-span, delayed reproduction, high parental investment in offspring (Pianka, 1970). The strategies also predict different responses towards environmental changes and biotic interactions, being r-strategists more prone to density-dependent mortality and unpredictable population changes, while K-strategists are favored in stable environments, keeping constant population sizes, and competing better for resources (MacArthur & Wilson, 1967; Pianka, 1970). Thus, GD would correlate with life-history traits that influence how population size changes over time. The explanation is that r-strategists can maintain larger N_e or recover faster after a population size fluctuation than K-strategists. It is important to notice that r-strategists can be more affected by momentaneous environmental changes (i.e., significant changes in current N_e) while K-strategists can avoid dramatic population size changes. However, these different strategies impact the historical N_e, which determines that r-strategists show high GD than K-strategists (Romiguier et al., 2014; Chen et al., 2017).

Our results show fecundity as a significant variable in just one dataset (dataset 1) and no influence on GD in the linear regression. But egg diameter (a proxy for propagule size) was selected as an important variable in two datasets (datasets 1 and 6, but significant only in the latter case) and showed significant results in the single variable analysis. This result is unexpected as fecundity and egg diameter are usually correlated and follow the r/K spectrum in crabs; hence they should result in the same outcomes (Hines, 1982). Some studies have shown that propagule size and fecundity strongly predict GD across many animals (Romiguier et al., 2014), but in some cases, their effect is taxon dependent (Chen et al., 2017). For instance, fishes were found to show a positive association between fecundity and GD (Martinez et al., 2018), or a negative association with egg diameter and no relationship with

fecundity (Mitton & Lewis Jr, 1989), while butterflies show no association for both variables (Mackintosh et al., 2019). Despite a general trend across animals, a closer look at some taxon may reveal different outcomes, indicating a strong taxon-dependent life history-GD association, such as the ones we found for crabs. A possible non-exclusive explanation for our results is the sweepstakes reproductive success (SRS) hypothesis (Hedgecock & Pudovkin, 2011). Crabs, like other marine species, produce many eggs that develop into larvae that are released in the water column. The selective pressures acting on the larvae can lead to variance in the individual reproduction success because most of the offspring die before metamorphosing to adults, thus influencing their contribution to the future genetic pool despite the larvae genetic pool. The SRS predicts that high fecundity might not be related to high N_e due to random processes acting on which set of larvae will contribute to the next generation, disrupting the association between fecundity and GD. Interestingly, Fratini et al. (2016) show a negative association between fecundity and GD for seven mangrove crabs, but this might result from their dataset, and when we combined more species data, we lost the association.

Even though we found that egg diameter (propagule size) is negatively correlated to GD, as expected, our association is much less powerful than previous studies analyzing many groups across the tree of life (Romiguier et al., 2014). Here, we assumed that the parent investment could be estimated by assessing egg diameter. Variation in egg investment can result in small feeding larvae (planktotrophic) or larvae with yolk reserves (lecithotrophic) (Vance et al., 1973). Many marine animals produce larvae during their life cycle, and in the case of crabs, most the juvenile and adult phases are benthic, and the larva is planktonic. Larvae may represent an adaptation to avoid competition between juveniles/adults and offspring, and it is also considered the animal's dispersal phase, having the potential to result in gene flow among populations, expanding species range, and even avoiding extinction (Pechenik, 1999). However, marine larvae are also constrained by local conditions like water circulation and selective forces acting on larvae but not adults, which potentially affects the overall association to GD, as suggested by the SRS. For instance, larvae can be subjected to local retention and accumulation, hampering dispersal (White et al., 2010), and the temperature in tropical regions might influence the

larval development duration by decreasing the dispersal time (Álvarez-Noriega et al., 2020). This means that egg size and fecundity might not be directly converted to the effects of r/K strategies on GD because these traits will not necessarily reflect the species susceptibility to environmental fluctuations (McEdward, 1997). Therefore, the propagule size/fecundity-GD association in species that show a larval phase is variable likely decreased compared to other taxa yet might show some contribution to the GD variation.

Traits related to dispersion ability (number of larval stages and pelagic larval duration - PLD) and habitat

We expected to find a correlation between the number of larval stages and PLD with GD, but this hypothesis was significant in just one analysis (dataset 1). Our expectations were based on the assumption that these traits could lead to connectivity among populations and, consequently, decrease the genetic drift effect and loss of GD (Shanks, 2009; Selkoe & Toonen, 2011; Faurby et al., 2012). Nevertheless, many studies do not find this association, adding complexity to understanding marine species dispersal (Weersing & Toonen, 2009; Butler et al., 2011; Iacchei et al., 2013; Timm et al., 2020). As we discussed before, this pattern may be explained by the difference between potential and realized dispersal due to ocean conditions that cause larvae retention and accumulation or quicken larval development in warmer regions (White et al., 2010; Hedgecock & Pudovkin, 2011; Álvarez-Noriega et al., 2020). Other factors might also affect dispersal, like larval behavior (Butler et al., 2011) and adults' behavior (Timm et al., 2020). Although we assumed an association between larval stages/PLD with connectivity, we did not estimate connectivity among crabs' populations.

Studies that estimate connectivity among populations and compared it across species found evidence that habitats vary at their connectivity level and that there might be habitat-GD associations (e.g., Harvey et al., 2017; Manel et al., 2020). Our results indicate GD variation across habitats; however, caution should be exercised because we could not precisely determine the pairwise differences, likely

due to the unbalanced comparisons. A visual inspection of our data indicates a considerable GD variation within habitats, except depth, which also seems to show lower GD. Indeed, depth can be associated with less differentiated populations and lower GD in the marine environment (Etter et al., 2005; García-Merchán et al., 2012; Selkoe et al., 2014). This might be explained by the stability of the deep-sea environment, which would favor specialization and refinement, generating low GD (Sanders, 1968; Bretsky & Lorenz, 1970). However, this hypothesis is questionable (McClain & Schlacher, 2015), and a compilation of marine invertebrate population genetic studies show their GD comparable to shallow-water species (Taylor & Roterman, 2017). We cannot draw precise conclusions from our data, but we are inclined to argue that "habitat" as a category is not the primary determinant variable, but a combination of local conditions and the associated disturbances species undergo together. For instance, marine turtles from the Atlantic expanded synchronously after the last glacial maximum, probably due to sea-level rise and fall during the Pleistocene, while Indo-Pacific lineages showed signs of stability (Reid et al., 2019). Consequently, Indo-Pacific lineages show higher GD than Atlantic lineages (Reid et al., 2019). Therefore, we believe demographic coalescent-based analyses investigating species that share the same habitat could provide a better framework to understand the habitat-GD association.

Limitations and future directions

We also have to take into consideration that we have used the COI (a mtDNA gene) to represent the species GD. The patterns of GD can vary significantly across the genome depending on the region and the forces acting upon it, having effects on N_e estimates (Charlesworth, 2009; Gossmann et al., 2011). Although many studies show an association between mtDNA and nuDNA variation (Mulligan et al., 2006; Nabholz et al., 2008a; Piganeau & Eyre-Walker, 2009), we did not test this assumption. As we mentioned before, the use of mtDNA is under debate, mainly due to its lack of recombination, susceptibility to positive selection, and selective sweeps, which can impact the GD (Ballard & Kreitman, 1995; Galtier et al., 2009). But considering the strong association of mtDNA GD and historical N_e, there

might be a correlation between crabs' mtDNA and nuDNA/N_e. We cannot confirm if the GD estimated using nuDNA would result in the same outcomes, but we also were constrained by the availability of genetic data and favored to maximize the number of species used. Also, we could not perform phylogenetic comparative methods because there is no robust phylogeny for Brachyura resolving deeper nodes or analyzing most of the diversity within crabs, which could compromise our inferences (Tsang et al., 2014; Ma et al., 2019; Tan et al., 2019; Wolfe et al., 2019). These limitations indicate future directions to be explored in crabs but also to be considered when studying other taxa.

For future studies, we suggest researchers deposit all sequences generated during their studies (not just unique haplotypes), always linking the sequences with a reference paper and using a standardized notation when submitting the sequence name. We recommend using the term "cytochrome c oxidase subunit I" to make it easier for public database searching algorithms to find all available sequences. Guaranteeing a standard submission name, short names like COI, coxI, CO1 can be used without interfering with the search. Therefore, studies that need to compile data will be benefited.

Conclusions

We provide insights on the association between ecological and life-history traits with crabs GD and contribute with taxon-specific results to the field investigating the determinants of GD. Our work takes advantage of publicly available COI sequence data to investigate the determinants of GD following the tendency of recent approaches (Miraldo et al., 2016; Manel et al., 2020; Theodoridis et al., 2020) but focusing on an invertebrate group. Using eight life history and demographic traits, we support the hypothesis of historical N_e being the most crucial factor determining GD, and we show less importance of life-history traits in contrast to other studies (Romiguier et al., 2014; Kort et al., 2021). The finding of strong N_e-GD association in crabs indicates that estimating mtDNA nucleotide diversity can be a first step assessing a species' health status. Effective population size and census population size are different concepts but usually show a proportional association (Frankham, 1995; Hauser & Carvalho, 2008).

Abundant fish species tend to show higher GD (McCusker & Bentzen, 2010), historical sampling of threatened species show a decrease in GD over time (Pichler & Baker, 2000; Tracy & Jamieson, 2011), and overfished species have been proved to show low GD in comparison to other species (Pinsky & Palumbi, 2014) indicating census population size may reflect N_e. Therefore, our results provide a baseline for comparison in future studies that aim to investigate crab's genetic diversity and their conservational status.

However, many other variables were not evaluated and could provide novel associations (e.g., species range size, age at maturity, latitude). Our outcomes might represent a taxa-dependent result, but we still lack studies investigating other taxa, especially marine invertebrates and specific groups among the crustaceans as well as, to expand our comparisons. Unfortunately, genetic, life history and ecological data may not be available for many invertebrate species. Further studies are warranted, and we encourage others to explore the influence of life-history traits in the GD of other groups to test which general trends represent, in fact, ecological and evolutionary rules.

GENERAL CONCLUSIONS

The chapters we have presented here had the goal of furthering the understanding on the patterns and processes of GD using crabs as model organisms. We begin (chapter 1) showing that an ecological trait (salinity tolerance) can be a best predictor of the species phylogeographic pattern than dispersal potential. We provide empirical evidence that co-distributed and closely related species can show contrasting GD spatial patterns based on a trait-based hypothesis elaborated *a priori*. However, when a species is tolerant to variation in abiotic conditions and it is not affected by biotic interactions, long pelagic dispersion can indeed guarantee the gene flow throughout the distributional range. We believe the assumption of PLD and lack of genetic structure should be considered as a null hypothesis and trait-based hypothesis could help us on a better understanding of the drivers of intraspecific GD spatial distribution. We further explored *C. ornatus* instraspecific diversification (chapter 2) by expanding our previous dataset and also employing a NGS approach (ddRAD-seq). Overall, mtDNA and ddRAD-seq show the same patterns but the latter gives us a better resolution. We propose an incipient scenario process driven by disruptive selection coupled with a permeable barrier (the Amazon-Orinoco plume), differences in effective population size, and oceanographic currents acting together in this complex scenario. Therefore, we show a possible consequence of the plume over the diversification of species distributed along the western Atlantic by investigating *C. ornatus.* We transitioned to a different question and explored GD spatial patterns (chapter 3) exploring the effects of latitude (LGDG), species range (CMH), a combination of both (CMH-LGDG), or species dispersal potential (MH). Investigating 14 species distributed along the western Atlantic, we show that there is an interspecific LGDG pattern (higher GD at lower latitudes) that is not explained by the evolutionary speed hypothesis (ESH). Alternatively, intraspecific GD spatial patterns varied across species and other hypotheses were evoked to explain our results. Finally, we addressed the question on the drivers of GD across crab species (chapter 4) by testing the effects of fecundity, body size, propagule size, number of larval stages, larval development time, historical effective population size, maximum longevity, and habitat. We compiled COI data for 150 species and confirm the primary influence of

historical N_e, although we show other traits having minor influence. Our results add taxon-based data to a growing body of literature exploring the determinants of GD across species and confirm and refute some previous trends. Additionally, showing the association between N_e and GD in crabs, we provide a baseline for future comparisons in conservation assessments.

In many ways, this dissertation only begins to hint at the mechanisms influencing GD in crabs. We provide novel results and interpretations for current questions in the ecology and evolutionary biology fields focusing on a group frequently overlooked. Although we answered some of our questions, many others arise from here. Future studies could look at similar patterns in other marine species (Chapter 1); be benefited by sampling in hybrid zones, exploring transcriptomes from the two divergent lineages, and performing controlled mating and salinity tolerance experiments (Chapter 2); exploring co-distributed species from other regions, including depth and longitude in the analysis, so as including other types of molecular markers (Chapter 3); performing phylogenetic comparative methods, increasing species sampling, employing different markers, investigating trends in phylogenetically fine-scale (Chapter 4). In conclusion, we gained great insight and substantially increased our knowledge of the GD patterns and processes operating in crabs and provided results that can be expanded to other taxa.

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., & Butlin, R.K. (2013). Hybridization and speciation. Journal of Evolutionary Biology, 26(2), 229-246.

Adams, R. I., & Hadly, E. A. (2013). Genetic diversity within vertebrate species is greater at lower latitudes. Evolutionary Ecology, 27(1), 133-143.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. Nature Reviews Genetics, 11(10), 697-709.

Allio, R., Donega, S., Galtier, N., & Nabholz, B. (2017). Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Molecular Biology and Evolution, 34(11), 2762-2772.

Álvarez-Campos, P., Giribet, G., & Riesgo, A. (2017). The *Syllis gracilis* species complex: a molecular approach to a difficult taxonomic problem (Annelida, Syllidae). Molecular Phylogenetics and Evolution, 109, 138-150.

Álvarez-Noriega, M., Burgess, S. C., Byers, J. E., Pringle, J. M., Wares, J. P., & Marshall, D. J. (2020). Global biogeography of marine dispersal potential. Nature Ecology & Evolution, 4(9), 1196-1203.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, 17(2), 81-92.

Anger, K., Queiroga, H., & Calado, R. (2015). Larval development and behaviour strategies in Brachyura. In: P. Castro; P. Davie; D. Guinot; F. Schram & C.V. Klein (Eds.), Treatise on Zoology – Anatomy, Taxonomy, Biology. The Crustacea, Volume 9 Part C. Brill, Leiden.

Antonovics, J. (2003). Toward community genomics? Ecology, 84(3), 598-601.

Arnold, M. L. (1992). Natural hybridization as an evolutionary process. Annual review of Ecology and Systematics, 23(1), 237-261.

Astorga, A., Fernández, M., Boschi, E. E., & Lagos, N. (2003). Two oceans, two taxa and one mode of development: latitudinal diversity patterns of South American crabs and test for possible causal processes. Ecology Letters, 6(5), 420-427.

Avise, J. C. (2000). Phylogeography: the history and formation of species. Harvard University Press.

Azizan, A., & Paradis, E. (2021). Patterns and drivers of genetic diversity among Felidae species. Biodiversity and Conservation, 30(2), 519-546.

Ballard, J. W. O., & Kreitman, M. (1995). Is mitochondrial DNA a strictly neutral marker? Trends in Ecology & Evolution, 10(12), 485-488.

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. Molecular Ecology, 13(4), 729-744.

Banks, S.C., Cary, G.J., Smith, A.L., Davies, I.D., Driscoll, D.A., Gill, A.M., Lindenmayer, D.B. & Peakall, R. (2013). How does ecological disturbance influence genetic diversity? Trends in Ecology & Evolution, 28(11), 670-679.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O., Swartz, B., Quental, T. B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., & Mersey, B. (2011). Has the Earth's sixth mass extinction already arrived? Nature, 471(7336), 51-57.

Barth, J.M., Berg, P.R., Jonsson, P.R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., Jakobsen, K.S., Johannesson, K., Jorde, P.E., Knutsen, H. & Moksnes, P.O. (2017). Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Molecular Ecology, 26(17), 4452-4466.

Barton, K. (2020). MuMIn: Multi-model inference. R package version 1.43.17. Available at: https://cran.r-project.org/web/packages/MuMIn/index.html

Battey, C. J., & Klicka, J. (2017). Cryptic speciation and gene flow in a migratory songbird species complex: Insights from the Red-eyed Vireo (*Vireo olivaceus*). Molecular Phylogenetics and Evolution, 113, 67-75.

Bazin, E., Glémin, S., & Galtier, N. (2006). Population size does not influence mitochondrial genetic diversity in animals. Science, 312(5773), 570-572.

Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics, 22(3), 341-345.

Beerli, P., & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences, 98(8), 4563-4568.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Molecular Ecology, 17(17), 3754-3774.

Beheregaray, L.B., Pfeiffer, L.V., Attard, C.R., Sandoval-Castillo, J., Domingos, F.M., Faulks, L.K., Gilligan, D.M., & Unmack, P.J. (2017). Genome-wide data delimits multiple climate-determined species ranges in a widespread Australian fish, the golden perch (*Macquaria ambigua*). Molecular Phylogenetics and Evolution, 111, 65-75.

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the american Lobster (*Homarus americanus*). Molecular Ecology, 24(13), 3299-3315.

Bertola, L.D., Boehm, J.T., Putman, N.F., Xue, A.T., Robinson, J.D., Harris, S., Baldwin, C.C., Overcast, I., & Hickerson, M.J. (2020). Asymmetrical gene flow in five co-distributed syngnathids explained by ocean currents and rafting propensity. Proceedings of the Royal Society B, 287(1926), 20200657.

Boldina, I., & Beninger, P. G. (2016). Strengthening statistical usage in marine ecology: Linear regression. Journal of Experimental Marine Biology and Ecology, 474, 81-91.

Bombonato, J.R., do Amaral, D.T., Silva, G.A.R., Khan, G., Moraes, E.M., da Silva Andrade, S.C., Eaton, D.A., Alonso, D.P., Ribolla, P.E.M., Taylor, N. & Zappi, D. (2020). The potential of genome-wide RAD sequences for resolving rapid radiations: a case study in Cactaceae. Molecular Phylogenetics and Evolution, 151, 106896.

Bowen, B.W., Shanker, K., Yasuda, N., Celia, M., Malay, M.C.M.D., von der Heyden, S., Paulay, G., Rocha, L.A., Selkoe, K.A., Barber, P.H. & Williams, S.T. (2014). Phylogeography unplugged: comparative surveys in the genomic era. Bulletin of Marine Science, 90(1), 13-46.

Bretsky, P. W., & Lorenz, D. M. (1970). An essay on genetic-adaptive strategies and mass extinctions. Geological Society of America Bulletin, 81(8), 2449-2456.

Briggs, J. C., & Bowen, B. W. (2013). Marine shelf habitat: biogeography and evolution. Journal of Biogeography, 40(6), 1023-1035.

Buranelli, R. C., & Mantelatto, F. L. (2017). Broad-ranging low genetic diversity among populations of the yellow finger marsh crab *Sesarma rectum* Randall, 1840 (Sesarmidae) revealed by DNA barcode. Crustaceana, 90(7-10), 845-864.

Buranelli, R. C., & Mantelatto, F. L. (2019). Comparative genetic differentiation study of three coexisting mangrove crabs in western Atlantic. Journal of Natural History, 53(47-48), 2883-2903.

Buranelli, R. C., Felder, D. L., & Mantelatto, F. L. (2019). Genetic diversity among populations of the Western Atlantic mangrove crab *Ucides cordatus* (Linnaeus, 1763) (Decapoda: Brachyura: Ocypodidae): evidence for panmixia and useful data for future management and conservation. Journal of Crustacean Biology, 39(4), 386-395.

Burney, C. W., & Brumfield, R. T. (2009). Ecology predicts levels of genetic differentiation in Neotropical birds. The American Naturalist, 174(3), 358-368.

Burnham, K. P. & Anderson, D. R (2002) Model selection and multimodel inference: a practical information-theoretic approach. 2nd ed. New York, Springer-Verlag.

Butler IV, M. J., Paris, C. B., Goldstein, J. S., Matsuda, H., & Cowen, R. K. (2011). Behavior constrains the dispersal of long-lived spiny lobster larvae. Marine Ecology Progress Series, 422, 223-237.

Catchen, J. M., Hohenlohe, P. A., Bernatchez, L., Funk, W. C., Andrews, K. R., & Allendorf, F. W. (2017). Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. Molecular Ecology Resources, 17(3), 362-365.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics, 10(3), 195-205.

Chaudhary, C., Saeedi, H., & Costello, M. J. (2016). Bimodality of latitudinal gradients in marine species richness. Trends in Ecology & Evolution, 31(9), 670-676.

Chaudhary, C., Saeedi, H., & Costello, M. J. (2017). Marine species richness is bimodal with latitude: a reply to Fernandez and Marques. Trends in Ecology & Evolution, 32(4), 234-237.

Chen, J., Glémin, S., & Lascoux, M. (2017). Genetic diversity and the efficacy of purifying selection across plant and animal species. Molecular Biology and Evolution, 34(6), 1417-1428.

Clement, M., Posada, D. C. K. A., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9(10), 1657-1659.

Collins, R. A., Wangensteen, O. S., O'Gorman, E. J., Mariani, S., Sims, D. W., & Genner, M. J. (2018). Persistence of environmental DNA in marine systems. Communications Biology, 1(1), 1-11.

Coyne, J. A., & Allen Orr, H. (1998). The evolutionary genetics of speciation. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353(1366), 287-305.

Cumer, T., Pouchon, C., Boyer, F., Yannic, G., Rioux, D., Bonin, A., & Capblancq, T. (2021). Double-digest RAD-sequencing: do pre-and post-sequencing protocol parameters impact biological results?. Molecular Genetics and Genomics, 296(2), 457-471.

Curtin, T. B. (1986). Physical observations in the plume region of the Amazon River during peak discharge—II. Water masses. Continental Shelf Research, 6(1-2), 53-71.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T. & McVean, G. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156-2158.

Davie, P.J.F., Guinot, D., & P.K.L. Ng. (2015). Systematics and classification of Brachyura, pp. 1049-1130. In: P. Castro; P. Davie; D. Guinot; F. Schram & C.V. Klein (Eds.), Treatise on Zoology – Anatomy, Taxonomy, Biology. The Crustacea, Volume 9 Part C. Brill, Leiden.

De Queiroz, K. (2007). Species concepts and species delimitation. Systematic Biology, 56(6), 879-886.

DeWoody, J. A., & Avise, J. C. (2000). Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fish Biology, 56(3), 461-473.

Díaz-Arce, N., & Rodríguez-Ezpeleta, N. (2019). Selecting RAD-Seq data analysis parameters for population genetics: the more the better?. Frontiers in Genetics, 10, 533.

Dion-Côté, A. M., & Barbash, D. A. (2017). Beyond speciation genes: an overview of genome stability in evolution and speciation. Current Opinion in Genetics & Development, 47, 17-23.

Dowle, E. J., Morgan-Richards, M., & Trewick, S. A. (2013). Molecular evolution and the latitudinal biodiversity gradient. Heredity, 110(6), 501-510.

Doyle, J. M., Hacking, C. C., Willoughby, J. R., Sundaram, M., & DeWoody, J. A. (2015). Mammalian genetic diversity as a function of habitat, body size, trophic class, and conservation status. Journal of Mammalogy, 96(3), 564-572.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources, 4(2), 359-361.

Eble, J. A., Toonen, R. J., & Bowen, B. W. (2009). Endemism and dispersal: comparative phylogeography of three surgeon fishes across the Hawaiian Archipelago. Marine Biology, 156(4), 689-698.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Molecular Ecology, 17(5), 1170-1188.

Edwards, S. V., Potter, S., Schmitt, C. J., Bragg, J. G., & Moritz, C. (2016). Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proceedings of the National Academy of Sciences, 113(29), 8025-8032.

Eizaguirre, C., Baltazar-Soares, M. (2014). Evolutionary conservation: evaluating the adaptive potential of species. Evolutionary Applications, 7, 963-967.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. Nature Reviews Genetics, 17(7), 422-433.

Eo, S. H., Doyle, J. M., & DeWoody, J. A. (2011). Genetic diversity in birds is associated with body mass and habitat type. Journal of Zoology, 283(3), 220-226.

Etter, R., Rex, M. A., Chase, M. R., & Quattro, J. M. (2005). Population differentiation decreases with depth in deep-sea bivalves. Evolution, 59(7), 1479-1491.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14(8), 2611-2620.

Faurby, S., & Barber, P. H. (2012). Theoretical limits to the correlation between pelagic larval duration and population genetic structure. Molecular Ecology, 21(14), 3419-3432.

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. Trends in Genetics, 28(7), 342-350.

Fernández, M., Astorga, A., Navarrete, S. A., Valdovinos, C., & Marquet, P. A. (2009). Deconstructing latitudinal species richness patterns in the ocean: does larval development hold the clue?. Ecology Letters, 12(7), 601-611.

Figueiredo, J. J. J. P., Hoorn, C., Van der Ven, P., & Soares, E. (2009). Late Miocene onset of the Amazon River and the Amazon deep-sea fan: Evidence from the Foz do Amazonas Basin. Geology, 37(7), 619-622.

Figuet, E., Nabholz, B., Bonneau, M., Mas Carrio, E., Nadachowska-Brzyska, K., Ellegren, H., & Galtier, N. (2016). Life history traits, protein evolution, and the nearly neutral theory in amniotes. Molecular Biology and Evolution, 33(6), 1517-1527.

Floeter, S.R., Rocha, L.A., Robertson, D.R., Joyeux, J.C., Smith-Vaniz, W.F., Wirtz, P., Edwards, A.J., Barreiros, J.P., Ferreira, C.E.L., Gasparini, J.L. & Brito, A. (2008). Atlantic reef fish biogeography and evolution. Journal of Biogeography, 35(1), 22-47.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics, 180(2), 977-993.

Fountain, E. D., Pauli, J. N., Reid, B. N., Palsbøll, P. J., & Peery, M. Z. (2016). Finding the right coverage: the impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. Molecular Ecology Resources, 16(4), 966-978.

Fox, H. E., Holtzman, J. L., Haisfield, K. M., McNally, C. G., Cid, G. A., Mascia, M. B., Parks, J.E, & Pomeroy, R. S. (2014). How are our MPAs doing? Challenges in assessing global patterns in marine protected area performance. Coastal Management, 42(3), 207-226.

François, O., Martins, H., Caye, K., & Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection. Molecular Ecology, 25(2), 454-469.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. Genetics Research, 66(2), 95-107.

Fu, Y. B. (2014). Genetic diversity analysis of highly incomplete SNP genotype data with imputations: an empirical assessment. G3: Genes, Genomes, Genetics, 4(5), 891-900.

Gaither, M. R., Toonen, R. J., Robertson, D. R., Planes, S., & Bowen, B. W. (2010). Genetic evaluation of marine biogeographical barriers: Perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). Journal of Biogeography, 37(1), 133-147.

Galtier, N., Nabholz, B., Glémin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular Ecology, 18(22), 4541-4550.

García-Merchán, V.H., Robainas-Barcia, A., Abelló, P., Macpherson, E., Palero, F., García-Rodríguez, M., de Sola, L.G. & Pascual, M. (2012). Phylogeographic patterns of decapod crustaceans at the Atlantic–Mediterranean transition. Molecular Phylogenetics and Evolution, 62(2), 664-672.

Garçon, D. P., Masui, D. C., Mantelatto, F. L., McNamara, J. C., Furriel, R. P. M., & Leone, F. A. (2007). K+ and NH4+ modulate gill (Na+, K+)-ATPase activity in the blue crab, *Callinectes ornatus*: fine tuning of ammonia excretion. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 147(1), 145-155.

Gargiulo, R., Kull, T., & Fay, M. F. (2020). Effective double-digest RAD sequencing and genotyping despite large genome size. Molecular Ecology Resources, 21, 1037-1055.

Gehara, M., Garda, A.A., Werneck, F.P., Oliveira, E.F., da Fonseca, E.M., Camurugi, F., Magalhães, F.D.M., Lanna, F.M., Sites Jr, J.W., Marques, R. & Silveira-Filho, R. (2017). Estimating synchronous demographic changes across populations using hABC and its application for a herpetological community from northeastern Brazil. Molecular Ecology, 26(18), 4756-4771.

Gillman, L. N., & Wright, S. D. (2014). Species richness and evolutionary speed: the influence of temperature, water and area. Journal of Biogeography, 41(1), 39-51.

Goodall-Copestake, W. P., Tarling, G. A., & Murphy, E. J. (2012). On the comparison of population-level estimates of haplotype and nucleotide diversity: a case study using the gene cox1 in animals. Heredity, 109(1), 50-56.

Gossmann, T. I., Woolfit, M., & Eyre-Walker, A. (2011). Quantifying the variation in the effective population size within a genome. Genetics, 189(4), 1389-1402.

Graham, C.F., Glenn, T.C., McArthur, A.G., Boreham, D.R., Kieran, T., Lance, S., Manzon, R.G., Martino, J.A., Pierson, T., Rogers, S.M., & Wilson, J.Y. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). Molecular Ecology Resources, 15(6), 1304-1315.

Gratton, P., Marta, S., Bocksberger, G., Winter, M., Keil, P., Trucchi, E., & Kühl, H. (2017). Which latitudinal gradients for genetic diversity?. Trends in Ecology & Evolution, 32(10), 724-726.

Gronau, I., Hubisz, M. J., Gulko, B., Danko, C. G., & Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nature Genetics, 43(10), 1031.

Guinot, D., Tavares, M., & Castro, P. (2013). Significance of the sexual openings and supplementary structures on the phylogeny of brachyuran crabs (Crustacea, Decapoda, Brachyura), with new nomina for higher-ranked podotreme taxa. Zootaxa, 3665(1), 1-414.

Gurgel, C. F. D., Camacho, O., Minne, A. J., Wernberg, T., & Coleman, M. A. (2020). Marine heatwave drives cryptic loss of genetic diversity in underwater forests. Current Biology, 30(7), 1199-1206.

Haefner Jr, P. A. (1990). Morphometry and size at maturity of *Callinectes ornatus* (Brachyura, Portunidae) in Bermuda. Bulletin of Marine Science, 46(2), 274-286.

Haponski, A. E., Lee, T., & Foighil, D. Ó. (2017). Moorean and Tahitian Partula tree snail survival after a mass extinction: New genomic insights using museum specimens. Molecular Phylogenetics and Evolution, 106, 151-157.

Hauser, L., & Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9(4), 333-362.

Hawkins, B. A. (2001). Ecology's oldest pattern?. Trends in Ecology & Evolution, 16(8), 470.

Hebert, P. D., Cywinska, A., Ball, S. L., & Dewaard, J. R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(1512), 313-321.

Hedgecock, D. (1986). Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates?. Bulletin of Marine Science, 39(2), 550-564.

Hedgecock, D., & Pudovkin, A. I. (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bulletin of Marine Science, 87(4), 971-1002.

Held, C. (2001). No evidence for slow-down of molecular substitution rates at subzero temperatures in Antarctic serolid isopods (Crustacea, Isopoda, Serolidae). Polar Biology, 24(7), 497-501.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359(1442), 183-195.

Hey, J., & Pinho, C. (2012). Population genetics and objectivity in species diagnosis. Evolution: International Journal of Organic Evolution, 66(5), 1413-1429.

Hillebrand, H. (2004a). On the generality of the latitudinal diversity gradient. The American Naturalist, 163(2), 192-211.

Hillebrand, H. (2004b). Strength, slope and variability of marine latitudinal gradients. Marine Ecology Progress Series, 273, 251-267.

Hines, A. H. (1982). Allometric constraints and variables of reproductive effort in brachyuran crabs. Marine Biology, 69(3), 309-320.

Hines, A. H. (1986). Larval patterns in the life histories of brachyuran crabs (Crustacea, Decapoda, Brachyura). Bulletin of Marine Science, 39(2), 444-466.

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Molecular Ecology Resources, 11, 117-122.

Hoorn, C., Paxton, C.G., Crampton, W.G., Burgess, P., Marshall, L.G., Lundberg, J.G., Räsänen, M.E., & Linna, A.M. (1996). Miocene deposits in the Amazonian foreland basin. Science, 122-125.

Hoorn, C., Wesselingh, F.P., Ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartín, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., & Jaramillo, C. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science, 330(6006), 927-931.

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Systematic Biology, 65(3), 357-365.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. Ecology Letters, 11(6), 609-623.

Hughes, L.C., Cardoso, Y.P., Sommer, J.A., Cifuentes, R., Cuello, M., Somoza, G.M., González-Castro, M., Malabarba, L.R., Cussac, V., Habit, E.M., & Betancur-R, R. (2020). Biogeography, habitat transitions and hybridization in a radiation of South American silverside fishes revealed by mitochondrial and genomic RAD data. Molecular Ecology, 29(4), 738-751.

Iacchei, M., Ben-Horin, T., Selkoe, K. A., Bird, C. E., García-Rodríguez, F. J., & Toonen, R. J. (2013). Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Molecular Ecology, 22(13), 3476-3494.

Ituarte, R. B., D'Anatro, A., Luppi, T. A., Ribeiro, P. D., Spivak, E. D., Iribarne, O. O., & Lessa, E. P. (2012). Population structure of the SW Atlantic estuarine crab *Neohelice granulata* throughout its range: a genetic and morphometric study. Estuaries and Coasts, 35(5), 1249-1260.

James, J., & Eyre-Walker, A. (2020). Mitochondrial DNA Sequence Diversity in Mammals: A Correlation between the Effective and Census Population Sizes. Genome Biology and Evolution, 12(12), 2441-2449.

Jansson, R., Rodríguez-Castañeda, G., & Harding, L. E. (2013). What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. Evolution, 67(6), 1741-1755.

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. Nature, 491(7424), 444-448.

Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. Trends in Ecology & Evolution, 19(2), 101-108.

Jombart, T., & Dray, S. (2010). adephylo: exploratory analyses for the phylogenetic comparative method. Bioinformatics, 26(15), 1-21.

Jormalainen, V., Danelli, M., Gagnon, K., Hillebrand, H., Rothäusler, E., Salminen, J. P., & Sjöroos, J. (2017). Genetic variation of a foundation rockweed species affects associated communities. Ecology, 98(11), 2940-2951.

Jost, J. A., Podolski, S. M., & Frederich, M. (2012). Enhancing thermal tolerance by eliminating the pejus range: a comparative study with three decapod crustaceans. Marine Ecology Progress Series, 444, 263-274.

Joyeux, J. C., Floeter, S. R., Ferreira, C. E. L., & Gasparini, J. L. (2001). Biogeography of tropical reef fishes: the South Atlantic puzzle. Journal of Biogeography, 28(7), 831-841.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods, 14(6), 587-589.

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.

Kelly, R. P., & Palumbi, S. R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. PloS one, 5(1), e8594.

Kiessling, W., Simpson, C., & Foote, M. (2010). Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. Science, 327(5962), 196-198.

Kinlock, N. L., Prowant, L., Herstoff, E. M., Foley, C. M., Akin-Fajiye, M., Bender, N., Umarani, M., Ryu, H.Y., Şen, B. & Gurevitch, J. (2018). Explaining global variation in the latitudinal diversity

gradient: Meta-analysis confirms known patterns and uncovers new ones. Global Ecology and Biogeography, 27(1), 125-141.

Knowlton, N. (1993). Sibling species in the sea. Annual Review of Ecology and Systematics, 24(1), 189-216.

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.

Kort, H., Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M., & Blanchet, S. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. Nature Communications, 12(1), 1-11.

Kulmuni, J., Butlin, R. K., Lucek, K., Savolainen, V., & Westram, A. M. (2020). Towards the completion of speciation: the evolution of reproductive isolation beyond the first barriers. Philosophical Transactions of the Royal Society B, 375:20190528.

Laurenzano, C., Costa, T. M., & Schubart, C. D. (2016). Contrasting patterns of clinal genetic diversity and potential colonization pathways in two species of western Atlantic fiddler crabs. PLoS One, 11(11), e0166518.

Laurenzano, C., Mantelatto, F. L., & Schubart, C. D. (2013). South American homogeneity versus Caribbean heterogeneity: population genetic structure of the western Atlantic fiddler crab *Uca rapax* (Brachyura, Ocypodidae). Journal of Experimental Marine Biology and Ecology, 449, 22-27.

Lawrence, E. R., & Fraser, D. J. (2020). Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity. Global Ecology and Biogeography, 29(5), 770-788.

Leasi, F., Andrade, S. C. D. S., & Norenburg, J. (2016). At least some meiofaunal species are not everywhere. Indication of geographic, ecological and geological barriers affecting the dispersion of species of *Ototyphlonemertes* (Nemertea, Hoplonemertea). Molecular Ecology, 25(6), 1381-1397.

Lefébure, T., Douady, C. J., Gouy, M., & Gibert, J. (2006). Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Molecular Phylogenetics and Evolution, 40(2), 435-447.

Leffler, E.M., Bullaughey, K., Matute, D.R., Meyer, W.K., Segurel, L., Venkat, A., Andolfatto, P. & Przeworski, M. (2012). Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biology, 10(9), e1001388.

Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6(9), 1110-1116.

Lewis, S. L., & Maslin, M. A. (2015). Defining the anthropocene. Nature, 519(7542), 171-180.

Liggins, L., Booth, D. J., Figueira, W. F., Treml, E. A., Tonk, L., Ridgway, T., Harris, D.A., & Riginos, C. (2015). Latitude-wide genetic patterns reveal historical effects and contrasting patterns of turnover and nestedness at the range peripheries of a tropical marine fish. Ecography, 38(12), 1212-1224.

Liggins, L., Gleeson, L., & Riginos, C. (2014). Evaluating edge-of-range genetic patterns for tropical echinoderms, *Acanthaster planci* and *Tripneustes gratilla*, of the Kermadec Islands, southwest Pacific. Bulletin of Marine Science, 90(1), 379-397.

Lin, G., Huang, Z., Wang, L., Chen, Z., Zhang, T., Gillman, L. N., & Zhao, F. (2019). Evolutionary rates of bumblebee genomes are faster at lower elevations. Molecular Biology and Evolution, 36(6), 1215-1219.

Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Molecular Ecology Resources, 19(3), 639-647.

Losos, J. B., & Glor, R. E. (2003). Phylogenetic comparative methods and the geography of speciation. Trends in Ecology & Evolution, 18(5), 220-227.

Lourenço, J. M., Glémin, S., Chiari, Y., & Galtier, N. (2013). The determinants of the molecular substitution process in turtles. Journal of Evolutionary Biology, 26(1), 38-50.

Lowerre-Barbieri, S. K., Ganias, K., Saborido-Rey, F., Murua, H., & Hunter, J. R. (2011). Reproductive timing in marine fishes: variability, temporal scales, and methods. Marine and Coastal Fisheries, 3(1), 71-91.

Lumpkin, R., & Johnson, G. C. (2013). Global ocean surface velocities from drifters: Mean, variance, El Niño–Southern Oscillation response, and seasonal cycle. Journal of Geophysical Research: Oceans, 118(6), 2992-3006.

Luna, F.D.O., Beaver, C.E., Nourisson, C., Bonde, R.K., Attademo, F.L., Miranda, A.V., Torres-Florez, J.P., De Sousa, G.P., Passavante, J.Z., & Hunter, M.E. (2021). Genetic connectivity of the West Indian manatee in the southern range and limited evidence of hybridization with Amazonian manatees. Frontiers in Marine Science, 7, 1089.

Luo, A., Lan, H., Ling, C., Zhang, A., Shi, L., Ho, S. Y., & Zhu, C. (2015). A simulation study of sample size for DNA barcoding. Ecology and Evolution, 5(24), 5869-5879.

Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources, 17(1), 67-77.

Lynch, M. (2010). Evolution of the mutation rate. Trends in Genetics, 26(8), 345-352.

Ma, K. Y., Qin, J., Lin, C. W., Chan, T. Y., Ng, P. K., Chu, K. H., & Tsang, L. M. (2019). Phylogenomic analyses of brachyuran crabs support early divergence of primary freshwater crabs. Molecular Phylogenetics and Evolution, 135, 62-66.

MacArthur, R. H. & Wilson, E.O. (1967). The theory of island biogeography. Princeton University Press, Princeton, New Jersey, USA.

Mackintosh, A., Laetsch, D. R., Hayward, A., Charlesworth, B., Waterfall, M., Vila, R., & Lohse, K. (2019). The determinants of genetic diversity in butterflies. Nature Communications, 10(1), 1-9.

Maier, P. A., Vandergast, A. G., Ostoja, S. M., Aguilar, A., & Bohonak, A. J. (2019). Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). Evolution, 73(12), 2476-2496.

Manel, S., Guerin, P. E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C., & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. Nature Communications, 11(1), 1-9.

Mantelatto, F.L. 2000. Allocation of the portunid crab *Callinectes ornatus* (Decapoda: Brachyura) in the Ubatuba Bay, northern coast of São Paulo State, Brazil. In: J.C. von Vaupel Klein & F. R. Schram (Eds.), The Biodiversity Crisis and Crustacea - Proceedings of the 4th International Crustacean Congress, Balkema, Rotterdam, Brookfield, The Netherlands. Crustacean Issues 12, 431-443.

Mantelatto, F. L., & Fransozo, A. (1999). Reproductive biology and moulting cycle of the crab *Callinectes ornatus* (Decapoda, Portunidae) from the Ubatuba region, São Paulo, Brazil. Crustaceana, 72(1), 63-76.

Mantelatto, F. L., & Fransozo, A. (2000). Brachyuran community in Ubatuba bay, northern coast of São Paulo state, Brazil. Journal of Shellfish Research, 19(2), 701-710.

Mantelatto, F. L., Tamburus, A. F., Magalhães, T., Buranelli, R. C., Terossi, M., Negri, M., Castilho, A.L., Costa, R.C., & Zara, F. J. (2020). Checklist of decapod crustaceans from the coast of the São Paulo state (Brazil) supported by integrative molecular and morphological data: III. Infraorder Brachyura Latreille, 1802. Zootaxa, 4872(1), 1-108.

Mantelatto, F.L., Robles, R., Biagi, R., & Felder, D.L. (2006) Molecular analysis of the taxonomic and distributional status for the hermit crab genera *Loxopagurus* Forest, 1964 and *Isocheles* Stimpson, 1858 (Decapoda, Anomura, Diogenidae). Zoosystema, 28(2), 495–506

Mantelatto, F.L., Terossi, M., Negri, M., Buranelli, R.C., Robles, R., Magalhães, T., Tamburus, A.F., Rossi, N. & Miyazaki, M.J. (2018) DNA sequence database as a tool to identify decapod crustaceans on the São Paulo coastline. Mitochondrial DNA Part A: DNA Mapping, Sequencing and Analysis, 29(5), 805–815.

Mardis, E. R. (2013). Next-generation sequencing platforms. Annual Review of Analytical Chemistry, 6, 287-303.

Marko, P. B., Hoffman, J. M., Emme, S. A., McGovern, T. M., Keever, C. C., & Nicole Cox, L. (2010). The 'Expansion–Contraction'model of Pleistocene biogeography: rocky shores suffer a sea change?. Molecular Ecology, 19(1), 146-169.

Marochi, M. Z., Masunari, S., & Schubart, C. D. (2017). Genetic and morphological differentiation of the semiterrestrial crab *Armases angustipes* (Brachyura: Sesarmidae) along the Brazilian coast. The Biological Bulletin, 232(1), 30-44.

Martin, A. P., & Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the molecular clock. Proceedings of the National Academy of Sciences, 90(9), 4087-4091.

Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J. & Jiggins, C.D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. Genome Research, 23(11), 1817-1828.

Martinez, A. S., Willoughby, J. R., & Christie, M. R. (2018). Genetic diversity in fishes is influenced by habitat type and life-history variation. Ecology and Evolution, 8(23), 12022-12031.

McClain, C. R., & Schlacher, T. A. (2015). On some hypotheses of diversity of animal life at great depths on the sea floor. Marine Ecology, 36(4), 849-872.

McCusker, M. R., & Bentzen, P. (2010). Positive relationships between genetic diversity and abundance in fishes. Molecular Ecology, 19(22), 4852-4862.

McEdward, L. R. (1997). Reproductive strategies of marine benthic invertebrates revisited: facultative feeding by planktotrophic larvae. The American Naturalist, 150(1), 48-72.

McKeown, N. J., & Shaw, P. W. (2008). Single paternity within broods of the brown crab Cancer pagurus: a highly fecund species with long-term sperm storage. Marine Ecology Progress Series, 368, 209-215.

Meirmans, P. G. (2020). genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. Molecular Ecology Resources, 20(4), 1126-1131.

Miglietta, M. P., Faucci, A., & Santini, F. (2011). Speciation in the sea: overview of the symposium and discussion of future directions. Integrative and Comparative Biology, 51(3), 449-455.

Miller, S., Dykes, D., & Polesky, H. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research, 16(3), 1215.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Molecular Biology and Evolution 37(5), 1530-1534.

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K.A & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. Science, 353(6307), 1532-1535.

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., Harrison, S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., & McCain, C. M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters, 10(4), 315-331.

Mitton, J. B., & Lewis Jr, W. M. (1989). Relationships between genetic variability and life-history features of bony fishes. Evolution, 43(8), 1712-1723.

Montgomery, M. E., Woodworth, L. M., Nurthen, R. K., Gilligan, D. M., Briscoe, D. A., & Frankham, R. (2000). Relationships between population size and loss of genetic diversity: comparisons of experimental results with theoretical predictions. Conservation Genetics, 1(1), 33-43.

Mori, G. M., Zucchi, M. I., & Souza, A. P. (2015). Multiple-geographic-scale genetic structure of two mangrove tree species: the roles of mating system, hybridization, limited dispersal and extrinsic factors. PloS ONE, 10(2), e0118710.

Moritz, C. T. E. D., Dowling, T. E., & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annual Review of Ecology and Systematics, 18(1), 269-292.

Mulligan, C. J., Kitchen, A., & Miyamoto, M. M. (2006). Comment on "population size does not influence mitochondrial genetic diversity in animals". Science, 314(5804), 1390-1390.

Mundry, R., & Nunn, C. L. (2009). Stepwise model fitting and statistical inference: turning noise into signal pollution. The American Naturalist, 173(1), 119-123.

Nabholz, B., Glémin, S., & Galtier, N. (2008b). Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Molecular Biology and Evolution, 25(1), 120-130.

Nabholz, B., Mauffrey, J. F., Bazin, E., Galtier, N., & Glemin, S. (2008a). Determination of mitochondrial genetic diversity in mammals. Genetics, 178(1), 351-361.

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. Molecular Ecology Resources, 11, 184-194.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. Molecular Ecology Resources, 17(6), 1136-1147.

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences, 76(10), 5269-5273.

Ng, P.K.L., Guinot, D., & Davie, P.J.R. (2008). Systema Brachyuororum: Part I. An annotated checklist of extant brachyuran crabs of the world. Raffles Bulletin of Zoology, 17, 1-286.

Nikolic, N., Liu, S., Jacobsen, M. W., Jónsson, B., Bernatchez, L., Gagnaire, P. A., & Hansen, M. M. (2020). Speciation history of European (*Anguilla anguilla*) and American eel (*A. rostrata*), analysed using genomic data. Molecular Ecology, 29(3), 565-577.

Nishikawa, K. S., Negri, M., & Mantelatto, F. L. (2021). Unexpected absence of population structure and high genetic diversity of the Western Atlantic hermit crab *Clibanarius antillensis* Stimpson, 1859 (Decapoda: Diogenidae) based on mitochondrial markers and morphological data. Diversity, 13(2), 56.

Norse, E. A. (1978). An experimental gradient analysis: hyposalinity as an" upstress" distributional determinant for Caribbean portunid crabs. The Biological Bulletin, 155(3), 586-598.

Nunes, F. L., Van Wormhoudt, A., Faroni-Perez, L., & Fournier, J. (2017). Phylogeography of the reef-building polychaetes of the genus *Phragmatopoma* in the western Atlantic Region. Journal of Biogeography, 44(7), 1612-1625.

Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. Evolution, 47(5), 1329-1341.

O'Grady, J. J., Brook, B. W., Reed, D. H., Ballou, J. D., Tonkyn, D. W., & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biological Conservation, 133(1), 42-51.

Ohta, T. (1992). The nearly neutral theory of molecular evolution. Annual Review of Ecology and Systematics, 23(1), 263-286.

Oppold, A. M., Pedrosa, J. A., Bálint, M., Diogo, J. B., Ilkova, J., Pestana, J. L., & Pfenninger, M. (2016). Support for the evolutionary speed hypothesis from intraspecific population genetic data in the non-biting midge *Chironomus riparius*. Proceedings of the Royal Society B: Biological Sciences, 283(1825), 20152413.

Orr, H. A., & Turelli, M. (2001). The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution, 55(6), 1085-1094.

Orton, M. G., May, J. A., Ly, W., Lee, D. J., & Adamowicz, S. J. (2019). Is molecular evolution faster in the tropics?. Heredity, 122(5), 513-524.

Pailler, K., Bourlès, B., & Gouriou, Y. (1999). The barrier layer in the western tropical Atlantic Ocean. Geophysical Research Letters, 26(14), 2069-2072.

Palumbi, S. R. (1992). Marine speciation on a small planet. Trends in Ecology & Evolution, 7(4), 114-118.

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics, 25(1), 547-572.

Pappalardo, P., & Fernández, M. (2014). Mode of larval development as a key factor to explain contrasting effects of temperature on species richness across oceans. Global Ecology and Biogeography, 23(1), 12-23.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35, 526-528.

Pedraza-Marrón, C.D.R., Silva, R., Deeds, J., Van Belleghem, S.M., Mastretta-Yanes, A., Domínguez-Domínguez, O., Rivero-Vega, R.A., Lutackas, L., Murie, D., Parkyn, D., & Bullock, L.H. (2019). Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. Proceedings of the Royal Society B, 286(1900), 20182924.

Pelc, R. A., Warner, R. R., & Gaines, S. D. (2009). Geographical patterns of genetic structure in marine species with contrasting life histories. Journal of Biogeography, 36(10), 1881-1890.

Peres, P. A., & Mantelatto, F. L. (2020). Salinity tolerance explains the contrasting phylogeographic patterns of two swimming crabs species along the tropical western Atlantic. Evolutionary Ecology, 34(4), 589-609.

Peres, P. A., Terossi, M., Iguchi, J., & Mantelatto, F. L. (2018). Can reproductive traits help to explain the coexistence of mud crabs *Panopeus* (Decapoda: Panopeidae)? A case of two sympatric species inhabiting an impacted mangrove area of Southern Brazil. Invertebrate Reproduction & Development, 62(3), 154-161.

Pertierra, L.R., Segovia, N.I., Noll, D., Martinez, P.A., Pliscoff, P., Barbosa, A., Aragón, P., Raya Rey, A., Pistorius, P., Trathan, P. & Polanowski, A. (2020). Cryptic speciation in gentoo penguins is driven by geographic isolation and regional marine conditions: Unforeseen vulnerabilities to global change. Diversity and Distributions, 26(8), 958-975.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS ONE, 7(5), e37135.

Phillips, J. D., Gillis, D. J., & Hanner, R. H. (2019). Incomplete estimates of genetic diversity within species: Implications for DNA barcoding. Ecology and Evolution, 9(5), 2996-3010.

Pianka, E. R. (1970). On r-and K-selection. The American Naturalist, 104(940), 592-597.

Pichler, F. B., & Baker, C. S. (2000). Loss of genetic diversity in the endemic Hector's dolphin due to fisheries-related mortality. Proceedings of the Royal Society of London. Series B: Biological Sciences, 267(1438), 97-102.

Piganeau, G., & Eyre-Walker, A. (2009). Evidence for variation in the effective population size of animal mitochondrial DNA. PloS ONE, 4(2), e4396.

Pinsky, M. L., & Palumbi, S. R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. Molecular Ecology, 23(1), 29-39.

Porter, T. M., & Hajibabaei, M. (2018). Over 2.5 million COI sequences in GenBank and growing. PloS ONE, 13(9), e0200177.

Potkamp, G., & Fransen, C. H. (2019). Speciation with gene flow in marine systems. Contributions to Zoology, 88(2), 133-172.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155(2), 945-959.

Puritz, J. B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., & Bird, C. E. (2014). Demystifying the RAD fad. Molecular Ecology, 23, 5937–5942.

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M., & Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. Nature, 559(7714), 392-395.

Reece, J. S., Bowen, B. W., Smith, D. G., & Larson, A. (2011). Comparative phylogeography of four Indo-Pacific moray eel species (Muraenidae) reveals comparable ocean-wide genetic connectivity despite five-fold differences in available adult habitat. Marine Ecology Progress Series, 437, 269-277.

Reid, B. N., Naro-Maciel, E., Hahn, A. T., FitzSimmons, N. N., & Gehara, M. (2019). Geography best explains global patterns of genetic diversity and postglacial co-expansion in marine turtles. Molecular Ecology, 28(14), 3358-3370.

Rivadeneira, M. M., & Poore, G. C. (2020). Latitudinal gradient of diversity of marine crustaceans: towards a synthesis. In: Thiel. M., & Poore, G. (Eds.), The natural history of the Crustacea: Evolution and Biogeography: Volume 8.

Rivadeneira, M. M., Thiel, M., González, E. R., & Haye, P. A. (2011). An inverse latitudinal gradient of diversity of peracarid crustaceans along the Pacific Coast of South America: out of the deep south. Global Ecology and Biogeography, 20(3), 437-448.

Rocha, L. A. (2003). Patterns of distribution and processes of speciation in Brazilian reef fishes. Journal of Biogeography, 30(8), 1161-1171.

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. Nature Protocols, 12(12), 2640-2659.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular Ecology, 28(21), 4737-4754.

Rohde, K. (1992). Latitudinal gradients in species diversity: the search for the primary cause. Oikos, 514-527.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N. & Loire, E. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature, 515(7526), 261-263.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution, 34(12), 3299-3302.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular Biology and Evolution, 34(12), 3299-3302.

Saenz-Agudelo, P., Dibattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. Molecular Ecology, 24(24), 6241-6255.

Sanders, H. L. (1968). Marine benthic diversity: a comparative study. The American Naturalist, 102(925), 243-282.

Sandström, A., Lundmark, C., Andersson, K., Johannesson, K., & Laikre, L. (2019). Understanding and bridging the conservation-genetics gap in marine conservation. Conservation Biology, 33(3), 725.

Schär, S., Vila, R., Petrović, A., Tomanović, Ž., Pierce, N. E., & Nash, D. R. (2017). Molecular substitution rate increases with latitude in butterflies: evidence for a trans-glacial latitudinal layering of populations? Ecography, 40(8), 930-935.

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution, 1(2), 103-113.

Schlaepfer, D. R., Braschler, B., Rusterholz, H. P., & Baur, B. (2018). Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: A meta-analysis. Ecosphere, 9(10), e02488.

Schlötterer, C. (2004). The evolution of molecular markers—just a matter of fashion?. Nature Reviews Genetics, *5*(1), 63-69.

Schluter, D., & Pennell, M. W. (2017). Speciation gradients and the distribution of biodiversity. Nature, 546(7656), 48-55.

Schubart, C. D., & Huber, M. G. J. (2006). Genetic comparisons of German populations of the stone crayfish, *Austropotamobius torrentium* (Crustacea: Astacidae). Bulletin Français de la Pêche et de la Pisciculture, (380-381), 1019-1028.

Schubart, C. D., Neigel, J. E., & Felder, D. L. (2000). Molecular phylogeny of mud crabs (Brachyura: Panopeidae) from the northwestern Atlantic and the role of morphological stasis and convergence. Marine Biology, 137(1), 11-18.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.P., Bank, C., Brännström, Å., & Brelsford, A. (2014). Genomics and the origin of species. Nature Reviews Genetics, 15(3), 176-192.

Selander, R. K., & Kaufman, D. W. (1973). Genic variability and strategies of adaptation in animals. Proceedings of the National Academy of Sciences, 70(6), 1875-1877.

Selkoe, K. A., & Toonen, R. J. (2011). Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Marine Ecology Progress Series, 436, 291-305.

Selkoe, K. A., Gaggiotti, O. E., Bowen, B. W., & Toonen, R. J. (2014). Emergent patterns of population genetic structure for a coral reef community. Molecular Ecology, 23(12), 3064-3079.

Shanks, A. L. (2009). Pelagic larval duration and dispersal distance revisited. The Biological Bulletin, 216(3), 373-385.

Sharifian, S., Kamrani, E., & Saeedi, H. (2020). Global biodiversity and biogeography of mangrove crabs: Temperature, the key driver of latitudinal gradients of species richness. Journal of Thermal Biology, 92, 102692.

Silva, D., Martins, K., Oliveira, J., da Silva, R., Sampaio, I., Schneider, H., & Gomes, G. (2018). Genetic differentiation in populations of lane snapper (*Lutjanus synagris* – Lutjanidae) from Western Atlantic as revealed by multilocus analysis. Fisheries Research, 198, 138-149.

Silva, J. M., Creer, S., Dos Santos, A., Costa, A. C., Cunha, M. R., Costa, F. O., & Carvalho, G. R. (2011). Systematic and evolutionary insights derived from mtDNA COI barcode diversity in the Decapoda (Crustacea: Malacostraca). PLoS ONE, 6(5), e19449.

Silveira, I. C. A. D., Schmidt, A. C. K., Campos, E. J. D., Godoi, S. S. D., & Ikeda, Y. (2000). A corrente do Brasil ao largo da costa leste brasileira. Revista Brasileira de Oceanografia, 48(2), 171-183.

Stankowski, S. (2013). Ecological speciation in an island snail: evidence for the parallel evolution of a novel ecotype and maintenance by ecologically dependent postzygotic isolation. Molecular Ecology, 22(10), 2726-2741.

Tamburus, A. F., & Mantelatto, F. L. (2016). Taxonomic and biogeographical status of three species of the spider crabs of the genus *Acanthonyx* Latreille, 1828 (Majoidea: Epialtidae) as determined by DNA barcoding and morphological analyses along the Western Atlantic. Zoological Studies, 55(23), 1-17.

Taylor, M. L., & Roterman, C. N. (2017). Invertebrate population genetics across Earth's largest habitat: The deep-sea floor. Molecular Ecology, 26(19), 4872-4896.

Terossi, M., Torati, L. S., Miranda, I., Scelzo, M. A., & Mantelatto, F. L. (2010). Comparative reproductive biology of two southwestern Atlantic populations of the hermit crab *Pagurus exilis* (Crustacea: Anomura: Paguridae). Marine Ecology, 31(4), 584-591.

Teske, P. R., Sandoval-Castillo, J., Golla, T. R., Emami-Khoyi, A., Tine, M., von der Heyden, S., & Beheregaray, L. B. (2019). Thermal selection as a driver of marine ecological speciation. Proceedings of the Royal Society B, 286(1896), 20182023.

Theodoridis, S., Fordham, D. A., Brown, S. C., Li, S., Rahbek, C., & Nogues-Bravo, D. (2020). Evolutionary history and past climate change shape the distribution of genetic diversity in terrestrial mammals. Nature Communications, 11(1), 1-11.

Thurman, C. L., Hopkins, M. J., Brase, A. L., & Shih, H. T. (2019). The unusual case of the widely distributed fiddler crab *Minuca rapax* (Smith, 1870) from the western Atlantic: an exemplary polytypic species. Invertebrate Systematics, 32(6), 1465-1490.

Timm, L. E., Isma, L. M., Johnston, M. W., & Bracken-Grissom, H. D. (2020). Comparative population genomics and biophysical modeling of shrimp migration in the Gulf of Mexico reveals current-mediated connectivity. Frontiers in Marine Science, 7, 19.

Timm, L., & Bracken-Grissom, H. D. (2015). The forest for the trees: evaluating molecular phylogenies with an emphasis on higher-level Decapoda. Journal of Crustacean Biology, 35(5), 577-592.

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. Nature, 466(7310), 1098-1101.

Titus, B. M., Blischak, P. D., & Daly, M. (2019). Genomic signatures of sympatric speciation with historical and contemporary gene flow in a tropical anthozoan (Hexacorallia: Actiniaria). Molecular Ecology, 28(15), 3572-3586.

Titus, B. M., Blischak, P. D., & Daly, M. (2019). Genomic signatures of sympatric speciation with historical and contemporary gene flow in a tropical anthozoan (Hexacorallia: Actiniaria). Molecular Ecology, 28(15), 3572-3586.

Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. Molecular Ecology, 21(16), 3907-3930.

Toms, J. A., Compton, J. S., Smale, M., & von der Heyden, S. (2014). Variation in palaeo-shorelines explains contemporary population genetic patterns of rocky shore species. Biology Letters, 10(6), 20140330.

Tourinho, J. L., Solé-Cava, A. M., & Lazoski, C. (2012). Cryptic species within the commercially most important lobster in the tropical Atlantic, the spiny lobster *Panulirus argus*. Marine Biology, 159(9), 1897-1906.

Tracy, L. N., & Jamieson, I. G. (2011). Historic DNA reveals contemporary population structure results from anthropogenic effects, not pre-fragmentation patterns. Conservation Genetics, 12(2), 517-526.

Tripp, E. A., Tsai, Y. H. E., Zhuang, Y., & Dexter, K. G. (2017). RAD seq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. Ecology and Evolution, 7(19), 7920-7936.

Trovant, B., Basso, N. G., Orensanz, J. M., Lessa, E. P., Dincao, F., & Ruzzante, D. E. (2016). Scorched mussels (*Brachidontes* spp., Bivalvia: Mytilidae) from the tropical and warm-temperate southwestern Atlantic: the role of the Amazon River in their speciation. Ecology and Evolution, 6(6), 1778-1798.

Tsang, L.M., Schubart, C.D., Ahyong, S.T., Lai, J.C., Au, E.Y., Chan, T.Y., Ng, P.K. & Chu, K.H. (2014). Evolutionary history of true crabs (Crustacea: Decapoda: Brachyura) and the origin of freshwater crabs. Molecular Biology and Evolution, 31(5), 1173-1187.

Valentine, J. W., & Jablonski, D. (2015). A twofold role for global energy gradients in marine biodiversity trends. Journal of Biogeography, 42(6), 997-1005.

Vance, R. R. (1973). On reproductive strategies in marine benthic invertebrates. The American Naturalist, 107(955), 339-352.

Vellend, M., & Geber, M. A. (2005). Connections between species diversity and genetic diversity. Ecology Letters, 8(7), 767-781.

Vellend, M., Lajoie, G., Bourret, A., Múrria, C., Kembel, S. W., & Garant, D. (2014). Drawing ecological inferences from coincident patterns of population-and community-level biodiversity. Molecular Ecology, 23(12), 2890-2901.

Vendrami, D.L., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M.S., Leese, F., McMinn, C., Moore, H., & Hoffman, J.I. (2017). RAD sequencing resolves fine-scale population

structure in a benthic invertebrate: implications for understanding phenotypic plasticity. Royal Society Open Science, 4(2), p.160548.

Villemereuil, P., Frichot, É., Bazin, É., François, O., & Gaggiotti, O. E. (2014). Genome scan methods against more complex models: when and how much should we trust them?. Molecular Ecology, 23(8), 2006-2019.

Vinagre, C., Leal, I., Mendonca, V., Madeira, D., Narciso, L., Diniz, M. S., & Flores, A. A. (2016). Vulnerability to climate warming and acclimation capacity of tropical and temperate coastal organisms. Ecological indicators, 62, 317-327.

Vogt, G. (2019). A compilation of longevity data in decapod crustaceans. Nauplius, 27, e2019011.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. Theoretical Population Biology, 7(2), 256-276.

Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. Marine Ecology Progress Series, 393, 1-12.

Wenger, S.J., Isaak, D.J., Luce, C.H., Neville, H.M., Fausch, K.D., Dunham, J.B., Dauwalter, D.C., Young, M.K., Elsner, M.M., Rieman, B.E. & Hamlet, A.F. (2011). Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change. Proceedings of the National Academy of Sciences, 108(34), 14175-14180.

White, J. W., Botsford, L. W., Hastings, A., & Largier, J. L. (2010). Population persistence in marine reserve networks: incorporating spatial heterogeneities in larval dispersal. Marine Ecology Progress Series, 398, 49-67.

Whitlock, R. (2014). Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: A meta-analysis. Journal of Ecology, 102(4), 857-872.

Willing, E. M., Dreyer, C., & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F ST do not necessarily require large sample sizes when using many SNP markers. PloS ONE, 7(8), e42649.

Willis, S. C., Hollenbeck, C. M., Puritz, J. B., Gold, J. R., & Portnoy, D. S. (2017). Haplotyping RAD loci: an efficient method to filter paralogs and account for physical linkage. Molecular Ecology Resources, 17(5), 955-965.

Wolfe, J. M., Luque, J., & Bracken-Grissom, H. D. (2021). How to become a crab: Phenotypic constraints on a recurring body plan. BioEssays, e2100020.

Wolfe, J.M., Breinholt, J.W., Crandall, K.A., Lemmon, A.R., Lemmon, E.M., Timm, L.E., Siddall, M.E. & Bracken-Grissom, H.D. (2019). A phylogenomic framework, evolutionary timeline and genomic resources for comparative studies of decapod crustaceans. Proceedings of the Royal Society B, 286(1901), 20190079.

Wooten, M. C., & Smith, M. H. (1985). Large mammals are genetically less variable? Evolution, 39(1), 210-212.

Xuereb, A., Benestan, L., Normandeau, É., Daigle, R. M., Curtis, J. M., Bernatchez, L., & Fortin, M. J. (2018). Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RAD seq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). Molecular Ecology, 27(10), 2347-2364.

Zupolini, L. L., Magalhães, T., Pileggi, L. G., & Mantelatto, F. L. (2017). Taxonomic revision of the speckled crabs, genus *Arenaeus* Dana, 1851 (Brachyura: Portunidae) based on morphological and molecular data. Zootaxa, 4273(3), 362-380.