Cecili Barrozo Mendes

Nemertopsis bivittata (Hoplonemertea), *Lineus sanguineus* (Heteronemertea) e *Perinereis ponteni* (Polychaeta): VALIDADE DAS ESPÉCIES, FLUXO GÊNICO E DIVERSIDADE GENÉTICA NA COSTA BRASILEIRA

Nemertopsis bivittata (Hoplonemertea), *Lineus sanguineus* (Heteronemertea) e *Perinereis ponteni* (Polychaeta): VALIDITY OF SPECIES, GENE FLOW AND GENETIC DIVERSITY ON BRAZILIAN COAST

São Paulo

2021

1

Cecili Barrozo Mendes

Nemertopsis bivittata (Hoplonemertea), *Lineus sanguineus* (Heteronemertea) e *Perinereis ponteni* (Polychaeta): VALIDADE DAS ESPÉCIES, FLUXO GÊNICO E DIVERSIDADE GENÉTICA NA COSTA BRASILEIRA

Nemertopsis bivittata (Hoplonemertea), *Lineus sanguineus* (Heteronemertea) e *Perinereis ponteni* (Polychaeta): VALIDITY OF SPECIES, GENE FLOW AND GENETIC DIVERSITY ON BRAZILIAN COAST

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Doutor em Biologia, na Área de Genética.

Orientadora: Sónia Cristina da Silva Andrade

Co-orientadora: Cinthya Simone Gomes Santos

São Paulo

2021

Ficha Catalográfica

Mendes, Cecili B.
    Nemertopsis bivittata (Hoplonemertea), *Lineus sanguineus* (Heteronemertea) e *Perinereis ponteni* (Polychaeta): VALIDADE DAS ESPÉCIES, FLUXO GÊNICO E DIVERSIDADE GENÉTICA NA COSTA BRASILEIRA / Cecili Barrozo Mendes; orientadora Sónia Cristina da Silva Andrade; Coorientadora Cinthya Simone Gomes do Santos – São Paulo, 2022.
    232 pp.

    Tese (Doutorado) – Instituto de Biociências da Universidade de São Paulo. Programa de Pós-Graduação em Biologia (genética).

    1. Genômica de paisagem. 2. Conectividade populacional. 3. Taxonomia. 4. Desenvolvimento larval. 5. Biologia marinha I. Andrade, Sónia Cristina da Silva, orient. II. Santos, Cinthya Simone Gomes, coorient. III. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva.

Comissão Julgadora:

_____
Prof(a). Dr(a).

_____
Prof(a). Dr(a).

_____
Prof(a). Dr(a).

_____
Prof(a). Dr(a).

_____
Profa. Dra. Sónia Cristina da Silva Andrade
Orientadora

Dedicatória

Às cientistas brasileiras que apesar de todas as dificuldades, sociais, financeiras e governamentais, continuam lutando para fazer do mundo um lugar melhor.

Epígrafe

But also when I am active *scientifically,* etc. – an activity which I can seldom perform in direct community with others – then my activity is *social,* because I perform it as a man. Not only is the material of my activity given to me as a social product (as is even the language in which the thinker is active): my own existence is social activity, and therefore that which I make of myself, I make of myself for society and with the consciousness of myself as a social being.

Karl Marx, *Reflections of a Young Man* (1835)

Agradecimentos

À minha família, especialmente à minha mãe Leonor e meu marido Paulo Pachelle pelo grande apoio. Minha mãe sempre me motiva a ser melhor e a sonhar grande. Meu marido é meu parceiro não apenas na vida pessoal, mas também na vida profissional, me auxiliando no campo em todas as horas e me fazendo ver a ciência por outro prisma. Também à minha filha canina, Julie e à minha filha felina, Martina. Elas são as princesas que me trazem alegria e conforto diários, sempre me fazendo ver o lado divertido da vida. Ainda às minhas primas, primos e tias queridas (Thamila, Susan, Clóvis, Leonardo, Larissa, Thamara, Arthur, Deborah, Thainá, Itamar, Lucy, Iris, Eduardo) pelo apoio emocional.

À minha orientadora Sónia Andrade, que me desafiou a ir mais longe do que imaginei que podia e me ensinou muito mais do que analisar dados e sequências. Sónia me fez crescer como profissional e me ajudou a entender que a carreira de cientista pode ser bastante difícil, mas é extremamente recompensadora. Ao meu vô científico (como ele mesmo se apelidou) Jon Norenburg, que foi meu primeiro mentor no trabalho com nemertíneos. Jon está sempre me apoiando e me ensinando sobre nemertíneos, taxonomia, ciência e vida em geral. À coorientadora Cinthya Santos pelos ensinamentos sobre a biologia de poliquetas e pela carinhosa acolhida em Niterói e em seu laboratório. Às amigas que o LDG me deu, Thainá Cortez e Tammy que me apoiaram nos momentos difíceis de laboratório, me acompanharam nos longos dias de trabalho e também nas comemorações dos objetivos atingidos. Aos outros colegas de laboratório, Gabriel Sonoda, Camilla, Dione, Rafael e Mylena que fizeram a vida de laboratório mais divertida e compartilharam das buscas por respostas aos mais estranhos erros de bioinformática. Aos mais novos companheiros, Marco, Flávia e Teresa que tanto me escutam falar nas reuniões de laboratório. Ainda aos outros amigos do porão, Tiago e Gi que também me emprestaram bastante seus ouvidos e me ajudaram a navegar mais facilmente

pelo mundo da pós graduação. Também aos funcionários do Departamento de Genética e Biologia Evolutiva, especialmente à Helenice por toda a ajuda com as burocracias.

Aos maravilhosos amigos que não só me acompanharam nas infindáveis coletas de campo (muitas vezes terminando sem encontrar um único nemertíneo), mas também proveram apoio logístico em outros estados, Tuane Ribeiro, Helena Mathews, Felipe Monteiro, Stephanie Pruefer, Thainá Cortez e claro, meu marido Pachelle. À minha querida amiga Gabrielle Rizzato, que conheci em uma disciplina do doutorado, e se tornou uma das minhas maiores confidentes e apoio durante todos os momentos sejam profissionais ou pessoais. Também aos amigos desde a graduação em biologia, Nayane, Anne, Marina, Marco, Raquel, Milu e Samuel que sempre me fizeram ver os outros lado biologia e perceber que eu não sou tão estranha assim.

Aos amigos que fiz durante meu período de doutorado sanduíche no Oregon (EUA), Svetlana e George que me receberam não só em seus laboratórios, mas também em sua família. Ambos me ensinaram incontáveis fatos sobre desenvolvimento larval e também sobre a vida no Oregon, me acolhendo em seus momentos familiares e me fazendo sentir tão querida. Também à Maureen, Christina, Mack, Nicole, Kara e Clara que me receberam de braços abertos em seu maravilhoso grupo de amigos, me fazendo sentir como parte dele desde o primeiro momento e nunca deixando me sentir sozinha. Vocês fizeram o Oregon parecer ser bem mais perto do Brasil.

Índice

General Introduction

The marine biodiversity and their taxonomy

The marine environment is the Earth's most extensive habitat. However, the marine biodiversity is not equally distributed, being the coastal region more densely populated and better known to science (APPELTANS et al., 2012; WEBB; BERGHE; O'DOR, 2010). The complete marine diversity, nonetheless, is still unknown. The average number of named marine species is ~410,000, of which ~390,000 are animal species, being ~330,000 of these invertebrate species ("WoRMS - World Register of Marine Species", 2021/11/22). According to taxonomists, the estimate number for the total of Eukaryote marine extant species is approximately four times this number, reaching 700 – 970 thousand species (APPELTANS et al., 2012).

Many of the unknown invertebrate species are part of cryptic species complexes that are not yet recognized or delimitated. Nonetheless, the cryptic species complexes have been more easily recognized due to the use of molecular tools that can differentiate species of similar morphology, which helped clarify the real distribution of many species believed to be cosmopolitan (JÖRGER et al., 2012; PUILLANDRE et al., 2012; SUNDBERG; KVIST; STRAND, 2016). This is especially important for animals of simpler morphology, and sometimes for animals with more complex morphology, for which the morphological characterization traditionally relays on histological preparations or other difficult to observe characters, such as wormlike animals as nemerteans, nematodes, peanut worms, among others (JOHNSON et al., 2016; JÖRGER et al., 2012; KAWAUCHI; GIRIBET, 2010; MENDES et al., 2018; STRAND; SUNDBERG, 2011; SUNDBERG et al., 2009, 2016; SUNDBERG; KVIST; STRAND, 2016). However, genetic data information use is better

suited when coupled with morphological, ecological, environmental and behavioral data. In addition, the utilization of a single gene in species delimitation studies might bias the conclusions, since the gene history might not be the same as the species history (KNOWLES; CARSTENS, 2007; MURPHY et al., 2015). One of the important causes for this bias is the Incomplete Lineage Sorting, when the lineages do not coalesce into a recent ancestor due to, e.g., a rapid diversification process (MADDISON, 1997; PAMILO; NEI, 1988). This can have meaningful implications to phylogenetic studies, and the most common way to overcame it is to add more or longer markers in the analysis (HODEL et al., 2020; NICHOLS, 2001; WANG et al., 2018). The use of new genetic tools, such as the fairly recent genomic approaches to understand the population structure can ensure that the species delimitation will not be artificial or influenced by a single locus unique or divergent history.

Gene flow in animal marine populations

The gene flow among marine populations is usually directly linked to the species reproduction mode. Usually, species with direct development will have more structured populations than species with indirect development. Among the indirect developers, usually species with planktotrophic larva (i.e. larva that feeds in the plankton, taking a longer time to settle) will have better connectivity than species with lecithotrophic larva (i.e. larva that do not feed in the plankton, using only the yolk already present in the egg, staying for shorter times in the water column until settle) (BOHONAK, 1999; HELLBERG et al., 2002; KINLAN; GAINES; LESTER, 2005; PALUMBI, 1994; RIESGO; TABOADA; AVILA, 2015). In addition to the time that the larvae stay in the plankton, their behaviour in the water column also affects the distance they can reach in this period. Most larvae cannot change their horizontal position, but they can travel vertically in the water column, and some do daily migrations to escape predators and UV-light. These movements can change the speed

and direction of their dispersion due to the different ocean currents in the different depths (BECKER et al., 2007; COWEN; SPONAUGLE, 2009; LEVIN, 2006).

Other factors like adult dispersion, local population extinction and geographic barriers also have an important influence in the population connectivity and might explain possible pattern deviations (ANDRADE; NORENBURG; SOLFERINI, 2011; FERNÁNDEZ et al., 2015). Another important factor to the population connectivity is the natural selection, which can act very differently in distinct environments, leading to differentiation between populations and limiting the gene flow. Such process can happen even among species with high dispersion rates if the local selection is strong enough (NOSIL; FUNK; ORTIZ-BARRIENTOS, 2009; VIA, 2002). Therefore, the study of patterns of connectivity between populations must take into account all possible influences, like the species biology, recent and ancient environmental changes, and possible geographic barriers (BANKS et al., 2007; BERNATCHEZ et al., 2019; SELKOE; TOONEN, 2011).

With the advent of DNA amplification and sequencing, these studies could have a better resolution and, especially among marine populations, unravel patterns of connectivity and differentiation rather unexpected (e.g. AMENT-VELÁSQUEZ et al., 2016; ANDRADE; NORENBURG; SOLFERINI, 2011; BOWEN et al., 2006; CARD et al., 2016; MENDES et al., 2018; NUNES et al., 2021; SELKOE; TOONEN, 2011). Most population genetic and phylogeography studies use mitochondrial markers, as they are haploid, non-recombinant and probably neutral (HELLBERG et al., 2002). However, some questions of fine scale structure still could not be resolved with the use of only mitochondrial regions. Multiple genes, both nuclear and mitochondrial, microsatellites and especially SNPs (Single Nucleotide Polymorphisms) can also be used for population studies, showing a better resolution in most cases. The cheapening of genomic scale sequencing enabled a broader use of SNPs, and a

new strategy combining SNPs and environmental features was developed (LIGGINS; TREML; RIGINOS, 2019). This approach is called landscape genomics when used to understand the role of the environment and geography on the terrestrial populations connectivity and seascape genomics, when focused on marine populations. In both cases, the patterns of connectivity and differentiation among populations are studied taking the environment effect into account, which allows for the discovery of possible loci under selection due to environment adaptation (SELKOE et al., 2016). The seascape genetics approach has allowed researchers to better understand connectivity patterns and evolution of well-known species, like abalones, oysters, lobsters and clown fish (BERNATCHEZ et al., 2019; SAENZ-AGUDELO et al., 2015; SANDOVAL-CASTILLO et al., 2018; SINGH et al., 2018). However, the application of seascape genetics approach in non-commercial and non-model species is still scarce.

The studied species

To better understand how the reproductive mode can shape population connectivity and local adaptation, one can compare species with different reproductive modes, but inhabiting the same environment. Along the Brazilian coast oyster and barnacle beds are quite common, and many different animal groups inhabit this environment. Among those, we have nemerteans and annelid polychaetes. Nemerteans are mostly benthic marine worms that possess a muscular proboscis housed in a coelomic cavity, the rhynchocoel, which characterizes the phylum. Three classes compose the phylum Nemertea, Palaeonemertea, Pilidiophora and Hoplonemertea (STRAND et al., 2019). They are morphologically distinguished by the position of the mouth and armature of the proboscis. Among palaeonemerteans and pilidiophorans the mouth opens ventrally, separately from the proboscis pore, and the proboscis is unarmed. Among hoplonemerteans the mouth shares an

opening with the proboscis pore in the tip of the head, and the proboscis is armed with one or many chitinous stylets (ANDRADE et al., 2012; THOLLESSON; NORENBURG, 2003). The larval development is also different between these three classes; most pilidiophorans have a hat-shaped planktotrophic larva that can stay up to several weeks in the plankton, feeding upon microalgae; the palaeonemerteans also present a planktotrophic larva, but these are planuliform and feed upon other invertebrate larvae; the hoplonemerteans also have a planuliform larva, but their development is only known for a few lecithotrophic species (BIRD; VON DASSOW; MASLAKOVA, 2014; MASLAKOVA; HIEBERT, 2014; MASLAKOVA; VON DÖHREN, 2009). However, previous population genetic studies on Hoplonemertea species indicate high connectivity between populations hundreds of kilometers apart, indicating that these species might have a long lived larva (ANDRADE; NORENBURG; SOLFERINI, 2011; LEASI; ANDRADE; NORENBURG, 2016; MENDES et al., 2018; TULCHINSKY; NORENBURG; TURBEVILLE, 2012). In addition, these species have eggs smaller than 100µm, a strong indicative of a planktotrophic larva (MASLAKOVA; HIEBERT, 2014).

Polychaeta is paraphyletic class of the phylum Annelida characterized by the presence of cephalic and locomotor appendages, the parapodia (AMARAL; RIZZO; ARRUDA, 2005; ANDRADE et al., 2015; ZRZAVÝ et al., 2009). Based on current knowledge Annelida is recognized as the two major clades of Errantia (polychaete families) and Sedentaria (that includes the traditional Clitellata,and Echiura, and polychete families). The two clades together form Pleistoannelida. Besides, there are more five basal branching lineages Sipuncula, Amphinomida, Chaetopteridae, Magelonidae, and Oweniidae (WEIGERT & BLEIDORN, 2016).

They have a well demarcated head, with a prostomium and peristomium, and sensorial and feeding appendages, as antennae, tentacular cirri and palps, may be present. Along the body they present parapodia and chaetae, that may show variation in shape, development, and number, respectively. They can also present gills, retractable or not (AMARAL; NONATO, 1996; ROUSE & PLEIJEL, 2001). These animals are mainly free-living, and most species are benthic and solitary. Other few species are planktonic and use their parapodia as paddles (AMARAL; NONATO, 1996; AMARAL; RIZZO; ARRUDA, 2005). The benthic polychaetes can live in many environments, like algae mats, oyster, mussel and barnacle beds, beach rocks, from the intertidal to the deep sea, etc. These species can be omnivorous, carnivores, scavengers, suspensivorous or herbivores (FAUCHALD; JUMARS, 1979; JUMARS et al., 2015). Some species are an important part of the diet of fishes and crustaceans of commercial value (PAIVA, 2006). Polychaete species present both planktotrophic and lecithotrophic larvae, and usually the larval types are consistent within each family, with a few exceptions (AMARAL; NONATO, 1996; PAIVA, 2006; ROUSE, 2000; ROUSE & PLEIJEL, 2001).

Taking that into consideration, the polychaete species *Perinereis ponteni* Kinberg, 1865, and the nemerteans *Lineus sanguineus* (Rathke, 1799) and *Nemertopsis bivittata* (Delle Chiaje, 1841) seem to be a good choice for a comparison study aiming to understand how the environmental conditions along the Brazilian coast influence the intertidal invertebrate populations. All three species live in the intertidal zone, usually among *Brachidontes* sp. mats, barnacle and oyster beds (Fig. 1), from Northeast to South Brazil. The polychaete *P. ponteni* (Fig. 2A) is an important generalist in these communities. This species present epitoky with a planktonic stage and might have a lecithotrophic larva, as most Nereididae do (BAKKEN et al., 2018; ROUSE, 2000). The nemerteans *L. sanguineus* (Fig. 2B) and *N. bivittata* (Fig. 2C) are predators. *Lineus sanguineus* usually preys upon syllid polychaetes

(CAPLINS; TURBEVILLE, 2011; RUNNELS, 2013), while *N. bivittata* prefers acorn barnacles (CAPLINS; NORENBURG; TURBEVILLE, 2012). In both cases, they use toxins to paralyze their prey, and can feed upon other invertebrates. Their reproduction strategies, however, are quite dissimilar, as *L. sanguineus* is a Pilidiophoran, which presents both sexual and asexual reproduction strategies. Some authors observed gonads under development, but gamete spawn was never observed, while asexual reproduction through fission is commonly observed (COE, 1899; MORETTO; BRANCATO, 1997; RUNNELS, 2013). The reproduction of the hoplonemertean *N. bivittata* is also not completely known, but their egg size smaller than 100µm. From the three studied species, this is the one with less information about their development available, since we do not have information on the development of any close related species.
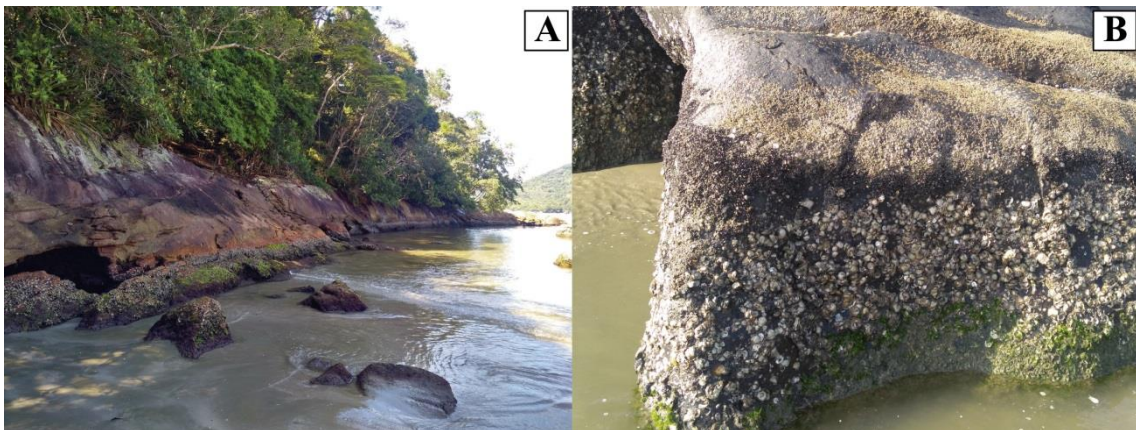
Figure 1. The environment where the species were collected. A: general view; B: zoom in the cluster of barnacles and oysters in the lower half, and mussels in the upper half.
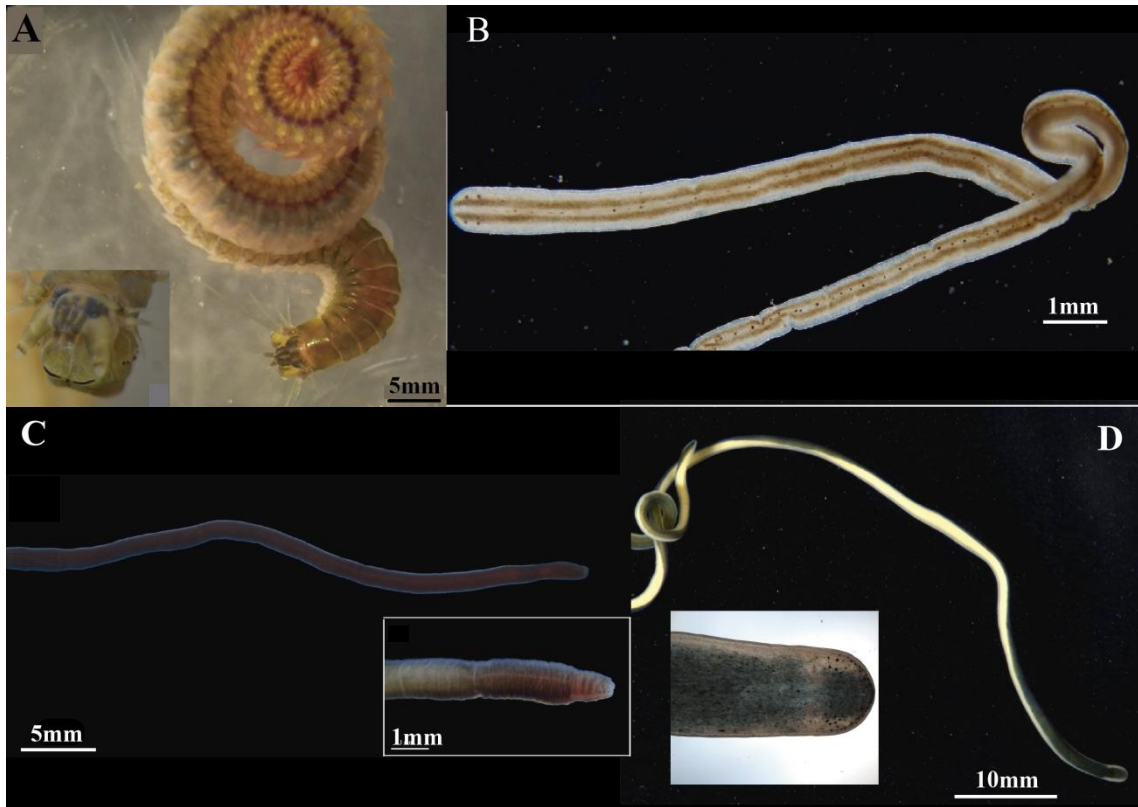
Figure 2. The four studied species. A: *Perinereis ponteni*, detail of everted pharynx, dorsal view. B: *Nemertopsis bivittata*. C: *Lineus sanguineus*, detail of the head in lateral view. D: *Emplectonema viride*, detail of the head in dorsal view.

Nemertopsis bivittata is a species of the family Emplectonematidae, being closely related to *Emplectonema viride* Stimpson, 1857, a very common species in the Northeastern Pacific ocean, present from Alaska to California, living in a similar environment as *N. bivittata* in the Brazilian coast (MENDES et al., 2021). Therefore, investigating *E. viride* larval development, one can have better insights about the larval development of *N. bivittata*. The development of *E. viride* is also not well known, but different sized larvae were observed in plankton samples along the Oregon coast (HIEBERT, 2016; MENDES et al., 2021). These larvae were also occasionally observed feeding upon barnacle nauplii (VON DASSOW et al. in prep). Reproductive individuals of *E. viride* can be easily found among barnacle clusters in the coast of Oregon in winter months, and can reproduce in captivity, being a good choice as

a model to understand the hoplonemerteans larval development and behaviour. Our initial hypothesis is that *E. viride* and *N. bivittata* both have planktotrophic larvae, being indirect developers. We also expect that the populations of the indirect developers (*N. bivittata* and *P. ponteni*) will be better connected than the populations of the partial asexual one (*L. sanguineus*).

This thesis is divided in four chapters: the first encompass the larval development of *E. viride* with notes in its larval behaviour and feeding; the second presents the taxonomic revision of *N. bivittata* along the Brazilian coast, with descriptions of three new species; the third chapter deals with the seascape genomics of *P. ponteni* populations in Brazil; and the fourth chapter comprehends a comparative seascape genomics of the nemertean species present in this study.

Objectives

The main objective of this work is to understand if there are genomic regions that might be involved with the response to environment in species with different reproductive and development modes. In addition, to identify what are the most important environmental variables influencing such response, and if there are congruencies among these species. We compared three species, two with sexual reproduction and indirect development (being one with lecithotrophic larva, and another with possibly planktotrophic larva), and one with partially asexual reproduction. To do so, we have the following specific objectives:

1. Describe the hoplonemertean larval development using *Emplectonema viride* as a model;

2. Genetically characterize, evaluate and compare the populations of *Nemertopsis bivittata*, *Lineus sanguineus* and *Perinereis ponteni* along part of the Brazilian coast;

3. Estimate their migration patterns and demographic history along the Brazilian coast;

4. Detect and identify potential loci associated with environmental factors that might indicate local adaptation.

References

AMARAL, A. C. Z.; NONATO, E. F. Annelida polychaeta: características, glossário e chaves para famílias e gêneros da costa brasileira. Campinas, SP, Brasil: Editora da Unicamp, 1996.

AMARAL, A.; RIZZO, A.; ARRUDA, E. Manual de Identificação dos Invertebrados Marinhos da Região Sudeste-Sul do Brasil. 1. ed. São Paulo, SP: EdUSP, 2005. v. 1

AMENT-VELÁSQUEZ, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (Lineus, Nemertea): hybridization, polyploidy and the Meselson effect. Molecular Ecology, v. 25, n. 14, p. 3356–3369, jul. 2016.

ANDRADE, S. C. S.; NORENBURG, J. L.; SOLFERINI, V. N. Worms without borders: genetic diversity patterns in four Brazilian *Ototyphlonemertes* species (Nemertea, Hoplonemertea). Marine Biology, v. 158, n. 9, p. 2109–2124, set. 2011.

ANDRADE, S. C. S. et al. Disentangling ribbon worm relationships: multi-locus analysis supports traditional classification of the phylum Nemertea. Cladistics, v. 28, n. 2, p. 141–159, abr. 2012.

ANDRADE, S. C. S. et al. Articulating "Archiannelids": Phylogenomics and Annelid Relationships, with Emphasis on Meiofaunal Taxa. Molecular Biology and Evolution, v. 32, n. 11, p. 2860–2875, nov. 2015.

APPELTANS, W. et al. The Magnitude of Global Marine Species Diversity. Current Biology, v. 22, n. 23, p. 2189–2202, dez. 2012.

BAKKEN, T. et al. Nereididae Blainville, 1818. In: Westheide, G. Purschke & M.Böggemann (Eds.), Handbook of zoology online (pp. 1–43). Berlin, Germany: DeGruyter.

BANKS, S. C. et al. Oceanic Variability and Coastal Topography Shape Genetic Structure in a Long-Dispersing Sea Urchin. Ecology, v. 88, n. 12, p. 3055–3064, 2007.

BECKER, B. J. et al. Complex larval connectivity patterns among marine invertebrate populations. Proceedings of the National Academy of Sciences, v. 104, n. 9, p. 3267–3272, fev. 2007.

BERNATCHEZ, S. et al. Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. Evolutionary Applications, v. 12, n. 3, p. 587–609, mar. 2019.

BIRD, A. M.; VON DASSOW, G.; MASLAKOVA, S. A. How the pilidium larva grows. EvoDevo, v. 5, n. 1, p. 13, 2014.

BOHONAK, A. J. Dispersal, Gene Flow, and Population Structure. The Quarterly Review of Biology, v. 74, n. 1, p. 21–45, mar. 1999.

BOUCHET, P. THE MAGNITUDE OF MARINE BIODIVERSITY. In: The Exploration of Marine Biodiversity: Scientific and Technological Challenges. Fundación BBVA, Madrid: C.M. Duarte, 2006. p. 31–62.

BOWEN, B. W. et al. Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): exploring links between pelagic larval duration and population connectivity. Marine Biology, v. 149, n. 4, p. 899–913, jul. 2006.

CAPLINS, S. A.; TURBEVILLE, J. M. The Occurrence of Ramphogordius sanguineus (Nemertea, Heteronemertea) in the Intertidal Zone of the Atlantic Coast of Virginia and New Observations on its Feeding Behavior. n. 38, p. 6, 2011.

CAPLINS, S.; NORENBURG, J. L.; TURBEVILLE, J. M. Molecular and morphological variation in the barnacle predator *Nemertopsis bivitatta* (Nemertea, Hoplonemertea). Integrative and Comparative Biology, n. 52, p. E24, 2012.

CARD, D. C. et al. Phylogeographic and population genetic analyses reveal multiple species of *Boa* and independent origins of insular dwarfism. Molecular Phylogenetics and Evolution, v. 102, p. 104–116, set. 2016.

COE, W. R. Notes on the Times of Breeding of Some Common New England Nemerteans. Science, v. 9, n. 214, p. 167–169, 1899.

CORTEZ, T. et al. Genome-wide assessment elucidates connectivity and the evolutionary history of the highly dispersive marine invertebrate *Littoraria flava* (Littorinidae: Gastropoda). Biological Journal of the Linnean Society, n. blab055, maio 2021.

COWEN, R. K.; SPONAUGLE, S. Larval Dispersal and Marine Population Connectivity. Annual Review of Marine Science, v. 1, n. 1, p. 443–466, 2009.

FAUCHALD, K.; JUMARS, P. A. The Diet of Worms: A Study of Polychaete Feeding Guilds. Oceanography and marine Biology annual review, p. 92, 1979.

FERNÁNDEZ, R. et al. Comparative phylogeography and population genetic structure of three widespread mollusc species in the Mediterranean and near Atlantic. Marine Ecology, v. 36, n. 3, p. 701–715, 2015.

FONTANETO, D.; FLOT, J.-F.; TANG, C. Q. Guidelines for DNA taxonomy, with a focus on the meiofauna. Marine Biodiversity, v. 45, n. 3, p. 433–451, set. 2015.

HELLBERG, M. E. et al. Genetic assessment of connectivity among marine populations. Bulletin of Marine Science, v. 70, n. 1, p. 18, 2002.

HERRERA, S.; SHANK, T. M. RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. Molecular Phylogenetics and Evolution, v. 100, p. 70–79, jul. 2016.

HIEBERT, T. C. NEW NEMERTEAN DIVERSITY DISCOVERED IN THE NORTHEAST PACIFIC, USING SURVEYS OF BOTH PLANKTONIC LARVAE AND BENTHIC ADULTS. University of Oregon, 2016.

HODEL, R. G. J. et al. When species trees disagree: an approach consistent with the coalescent that quantifies phylogenomic support for contentious relationships. Disponível em: <https://www.biorxiv.org/content/10.1101/2020.03.27.012237v1>. Acesso em: 26 ago. 2021.

JOHNSON, N. D. et al. Cryptic species in Pacific sipunculans (Sipuncula: Phascolosomatidae): east-west divergence between non-sister taxa. Zoologica Scripta, v. 45, n. 4, p. 455–463, 2016.

JÖRGER, K. M. et al. Barcoding against a paradox? Combined molecular species delineations reveal multiple cryptic lineages in elusive meiofaunal sea slugs. BMC Evolutionary Biology, v. 12, n. 1, p. 245, dez. 2012.

KAWAUCHI, G. Y.; GIRIBET, G. Are there true cosmopolitan sipunculan worms? A genetic variation study within Phascolosoma perlucens (Sipuncula, Phascolosomatidae). Marine Biology, v. 157, n. 7, p. 1417–1431, jul. 2010.

KINLAN, B. P.; GAINES, S. D.; LESTER, S. E. Propagule dispersal and the scales of marine community process: Marine dispersal scales. Diversity and Distributions, v. 11, n. 2, p. 139–148, mar. 2005.

KNOWLES, L. L.; CARSTENS, B. C. Delimiting Species without Monophyletic Gene Trees. Systematic Biology, v. 56, n. 6, p. 887–895, dez. 2007.

LEASI, F.; ANDRADE, S. C. DA S.; NORENBURG, J. 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, v. 25, n. 6, p. 1381–1397, 2016.

LEVIN, L. A. Recent progress in understanding larval dispersal: new directions and digressions. Integrative and Comparative Biology, v. 46, n. 3, p. 282–297, jun. 2006.

LIGGINS, L.; TREML, E. A.; RIGINOS, C. Seascape Genomics: Contextualizing Adaptive and Neutral Genomic Variation in the Ocean Environment. In: OLEKSIAK, M. F.; RAJORA, O. P. (Eds.). . Population Genomics: Marine Organisms. Population Genomics. Cham: Springer International Publishing, 2019. p. 171–218.

MADDISON, W. P. Gene Trees In Species Trees. Systematic Biology, v. 46, n. 3, p. 523–536, 1997.

MASLAKOVA, S. A.; VON DÖHREN, J. Larval Development with Transitory Epidermis in *Paranemertes peregrina* and Other Hoplonemerteans. The Biological Bulletin, v. 216, n. 3, p. 273–292, jun. 2009.

MASLAKOVA, S. A.; HIEBERT, T. C. From trochophore to pilidium and back again - a larva's journey. The International Journal of Developmental Biology, v. 58, n. 6-7–8, p. 585–591, 2014.

MATHEWS, L. M.; ANKER, A. Molecular phylogeny reveals extensive ancient and ongoing radiations in a snapping shrimp species complex (Crustacea, Alpheidae, *Alpheus armillatus*). Molecular Phylogenetics and Evolution, v. 50, n. 2, p. 268–281, fev. 2009.

MENDES, C. B. et al. Hidden diversity: Phylogeography of genus *Ototyphlonemertes* Diesing, 1863 (Ototyphlonemertidae: Hoplonemertea) reveals cryptic species and high diversity in Chilean populations. PLOS ONE, v. 13, n. 4, p. e0195833, abr. 2018.

MENDES, C. B. et al. Redescription of *Emplectonema viride* – a ubiquitous intertidal hoplonemertean found along the West Coast of North America. ZooKeys, v. 1031, p. 1–17, abr. 2021.

MORETTO, H. J. A.; BRANCATO, C. L. The ovaries of a fissiparous heteronemertean, Lineus bonaerensis, from Argentina. Hydrobiologia, v. 365, n. 1/3, p. 129–134, 1997.

MURPHY, N. P. et al. Species, ESUs or populations? Delimiting and describing morphologically cryptic diversity in Australian desert spring amphipods. Invertebrate Systematics, v. 29, n. 5, p. 457–467, out. 2015.

NICHOLS, R. Gene trees and species trees are not the same. Trends in Ecology & Evolution, v. 16, n. 7, p. 358–364, jul. 2001.

NOSIL, P.; FUNK, D. J.; ORTIZ-BARRIENTOS, D. Divergent selection and heterogeneous genomic divergence. Molecular Ecology, v. 18, n. 3, p. 375–402, 2009.

NUNES, F. L. D. et al. Looking for diversity in all the right places? Genetic diversity is highest in peripheral populations of the reef-building polychaete *Sabellaria alveolata*. Marine Biology, v. 168, n. 5, p. 63, maio 2021.

PAIVA, P. C. Capítulo 7 Filo Annelida Classe Polychaeta. Biodiversidade bentônica da região central da Zona Econômica Exclusiva brasileira, v. 18, p. 261–298, 2006.

PALUMBI, S. R. Genetic Divergence, Reproductive Isolation, and Marine Speciation. Annual review of ecology and systematics, v. 25, n. 1, p. 547-572, 1994.

PAMILO, P.; NEI, M. Relationships between Gene Trees and Species Trees. Molecular biology and evolution, v. 5, n. 5 p. 568–583, 1988.

PONS, J. et al. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. Systematic Biology, v. 55, n. 4, p. 595–609, 1 ago. 2006.

PUILLANDRE, N. et al. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology, v. 21, n. 8, p. 1864–1877, abr. 2012.

RIESGO, A.; TABOADA, S.; AVILA, C. Evolutionary patterns in Antarctic marine invertebrates: An update on molecular studies. Marine Genomics, v. 23, p. 1–13, out. 2015.

ROUSE, G. W. Polychaetes have evolved feeding larvae numerous times. Bulletin of MArine Science, v. 67, n. 1, p. 19, 2000.

RUNNELS, C. Phylogeography and Species Status of *Ramphogordius sanguineus*. Theses and Dissertations, 23 jul. 2013. Virginia Commonwealth University.

SAENZ-AGUDELO, P. et al. Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. Molecular Ecology, v. 24, n. 24, p. 6241–6255, dez. 2015.

SANDOVAL-CASTILLO, J. et al. Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. Molecular Ecology, v. 27, n. 7, p. 1603–1620, 2018.

SELKOE, K. et al. A decade of seascape genetics: contributions to basic and applied marine connectivity. Marine Ecology Progress Series, v. 554, p. 1–19, jul. 2016.

SELKOE, K.; TOONEN, R. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Marine Ecology Progress Series, v. 436, p. 291–305, ago. 2011.

SINGH, S. P. et al. Seascape genetics of the spiny lobster *Panulirus homarus* in the Western Indian Ocean: Understanding how oceanographic features shape the genetic structure of species with high larval dispersal potential. Ecology and Evolution, v. 8, n. 23, p. 12221–12237, 2018.

STRAND, M. et al. Nemertean taxonomy-Implementing changes in the higher ranks, dismissing Anopla and Enopla. Zoologica Scripta, v. 48, n. 1, p. 118–119, jan. 2019.

STRAND, M.; SUNDBERG, P. A DNA-based description of a new nemertean (phylum Nemertea) species. Marine Biology Research, v. 7, n. 1, p. 63–70, jan. 2011.

SUNDBERG, P. et al. Polymorphism hides cryptic species in *Oerstedia dorsalis* (Nemertea, Hoplonemertea): CRYPTIC SPECIES IN O. DORSALIS. Biological Journal of the Linnean Society, v. 98, n. 3, p. 556–567, out. 2009.

SUNDBERG, P. et al. The future of nemertean taxonomy (phylum Nemertea) - a proposal. Zoologica Scripta, v. 45, n. 6, p. 579–582, nov. 2016.

SUNDBERG, P.; KVIST, S.; STRAND, M. Evaluating the Utility of Single-Locus DNA Barcoding for the Identification of Ribbon Worms (Phylum Nemertea). PLOS ONE, v. 11, n. 5, p. e0155541, maio 2016.

TABOADA, S.; PÉREZ-PORTELA, R. Contrasted phylogeographic patterns on mitochondrial DNA of shallow and deep brittle stars across the Atlantic-Mediterranean area. Scientific Reports, v. 6, n. 1, p. 32425, set. 2016.

THOLLESSON, M.; NORENBURG, J. L. Ribbon worm relationships: a phylogeny of the phylum Nemertea. Proceedings of the Royal Society of London. Series B: Biological Sciences, v. 270, n. 1513, p. 407–415, fev. 2003.

TULCHINSKY, A. Y.; NORENBURG, J. L.; TURBEVILLE, J. M. Phylogeography of the marine interstitial nemertean *Ototyphlonemertes parmula* (Nemertea, Hoplonemertea) reveals cryptic diversity and high dispersal potential. Marine Biology, v. 159, n. 3, p. 661–674, mar. 2012.

VIA, S. The Ecological Genetics of Speciation. The American Naturalist, v. 159, n. S3, p. S1–S7, mar. 2002.

WANG, K. et al. Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. Communications Biology, v. 1, n. 1, p. 1–9, out. 2018.

WEBB, T. J.; BERGHE, E. V.; O'DOR, R. Biodiversity's Big Wet Secret: The Global Distribution of Marine Biological Records Reveals Chronic Under-Exploration of the Deep Pelagic Ocean. PLOS ONE, v. 5, n. 8, p. e10223, ago. 2010.

WoRMS - World Register of Marine Species. Disponível em: <http://www.marinespecies.org/aphia.php?p=stats>. Acesso em: 22 nov. 2021.

ZRZAVÝ, J. et al. Phylogeny of Annelida (Lophotrochozoa): total-evidence analysis of morphology and six genes. BMC Evolutionary Biology, v. 9, n. 1, p. 189, ago. 2009.

Chapter 1 . Complete development and feeding larva of *Emplectonema viride* (Emplectonematidae, Monostilifera) under laboratory conditions: from egg to egg

Cecili B. Mendes, George von Dassow, Sónia C. S. Andrade & Svetlana Maslakova

Abstract

The early stages of hoplonemertean development are well known in a variety of species, but larval development through settlement is only described for a handful of direct-developing species. All hoplonemertean species were presumed to have lecithotrophic direct development until the recent discovery of macrophagous planktotrophy in six species, including *Emplectonema viride*, a common intertidal hoplonemertean whose complete development from egg through planktotrophic larva to reproductive adult we describe here. We achieved this by feeding the larvae with planktonic crustaceans. They primarily ate barnacle nauplii and cyprids, but also fed upon calanoid copepods. The larvae grew for about 120 days in lab culture with abundant food, before settling as juveniles and starting to feed upon adult barnacles, with first juvenile appearing at day 90. Settlement is accompanied by a subtle but definite metamorphosis, which includes shortening of epidermal ciliation and loss of the caudal cirrus, in addition to behavioral changes. Larvae are positively phototactic, whereas juveniles are negatively phototactic. Such long larval stage has direct influence on the dispersal potential, and hence the amount of gene flow between populations. Genetic population studies with hoplonemerteans reveal gene flow in excess of expectations for direct developers, and planktotrophy may be widespread within the class. Therefore, *E. viride* may represent a more typical hoplonemertean life cycle than previously-described direct developers.

Keywords: Nemertea; ribbon worms; larval development; planktotrophy.

Introduction

Planktotrophic larvae typically have a pelagic phase of weeks to months, depending on food quantity and quality, water temperature, and availability of settlement cues. Lecithotrophic larvae rely on yolk present in the egg, and although lecithotrophs may be present in the plankton, they do not feed. They tend to have shorter pelagic duration (hours to weeks). Many studies have linked the type of development (planktotrophic vs. lecithotrophic) to connectivity between adult populations of marine species (i.e. Levin *et al.* 1987; Bowen *et al.* 2006; Hellberg 2007). The longer the pelagic duration of the larval stage, the higher the potential for long-distance dispersal. The adult phase of many benthic marine invertebrates has low dispersal potential, while the larval phase is capable of much longer dispersal distances, and thus potentially maintains gene flow between populations (Cowen & Sponaugle 2009). During this stage, larvae may be passively carried horizontally by ocean currents, but they can also exhibit regular (e.g. daily) vertical migrations, which can affect their horizontal dispersal (Cowen & Sponaugle 2009). The pelagic duration of larval stage is determined both by environmental and biological characteristics and can vary even between closely related species (Levin *et al*. 1987; Cowen & Sponaugle 2009).

Nemerteans are mostly free-living marine predatory lophotrochozoan worms, related to mollusks and annelids (Struck and Fisse 2008; Podsiadlowski *et al.* 2009; Laumer *et al.* 2019). The phylum is characterized by having a dorsal cavity that houses a muscular proboscis used for both predation and defense. Both lecithotrophic and planktotrophic development are found within the phylum, and within each of the three classes: Pilidiophora, Hoplonemertea, and Palaeonemertea (Thollesson and Norenburg, 2003; Maslakova and Hiebert 2014; Andrade *et al*. 2014; Strand *et al.*

2019). The easily recognizable pilidiophoran planktotrophic larvae (pilidia) are typically shaped like a deerstalker cap, with a prominent apical tuft at the anterior end, and lobes (front and back "visors") and lappets ("ear flaps") at the posterior end. Pilidia feed on unicellular algae (von Dassow *et al.* 2013), and the juvenile develops inside the larval body from eight distinct rudiments (imaginal discs) over the course of weeks to months in the plankton. Once the juvenile is complete, the pilidium undergoes rapid and catastrophic metamorphosis, in which the juvenile erupts from and, typically, eats the larval body (Maslakova 2010). Hoplonemerteans and palaeonemerteans lack the pilidium. Instead, species with pelagic development have the so-called planuliform larva that superficially resembles the planula larva of some cnidarians — essentially, a planktonic juvenile (Maslakova *et al.,* 2004; Maslakova and Hiebert, 2014; Hiebert *et al*. 2010, Maslakova and von Doehren 2009).

While pelagic larvae of most palaeonemerteans are planktonic macrophagous predators, hoplonemertean larvae were thought to be lecithotrophic (Iwata 1960; Jägersten 1972; Stricker and Norenburg 2002; Maslakova 2010; Maslakova and Hiebert 2014). Indeed, some hoplonemertean species develop directly without feeding in the plankton (e.g. Maslakova and Malakhov 1999; Maslakova and von Döhren 2009). However, the presence in the plankton of larvae from the same species but with a broad spectrum of sizes (Maslakova and Hiebert 2014), or the strong gene flow among populations many kilometers apart (Andrade *et al.* 2011; Leasi *et al.* 2016; Mendes *et al.* 2018), are indications of species with a long-lived planktonic larva among hoplonemerteans. Indeed, recently von Dassow et al. (2021) published the first direct observation of planktotrophy through carnivorous feeding by the larvae of six hoplonemertean species, *Paranemertes californica*, *Paranemertes* sp., *Gurjanovella littoralis*, *Emplectonema viride*, *Carcinonemertes epialti*, and

Ototyphlonemertes sp*.,* showing images and sequence data of their prey, as well as their feeding mechanism.

Emplectonema viride Stimpson, 1857 is one of the most common free-living intertidal hoplonemerteans along temperate rocky shores in the NE Pacific, whose spawning has been observed in Oregon during most of the year. This species occurs along the Pacific Coast of North America (Griffin 1898; Gibson 1995; Mendes *et al.* 2021), lives in the upper-middle intertidal zone, and is found among barnacle and mussels in natural and anthropic environments, where it feeds upon acorn barnacles. The adults are long and slender with a dark green dorsal surface and cream-colored or pale yellow ventrally. The planuliform larva has a characteristic green color, and can be found in plankton samples mostly during winter months (Hiebert, 2016). Early developmentof *Emplectonema gracile* (Johnston, 1837), the sister species of *E. viride*, were described by Iwata (1960) through proboscis and stylet formation. However, because the hatchlings died without food after about two weeks, this study did not describe late larval development.

von Dassow *et al.* (2021) found that this species larva feed upon barnacle nauplii and cyprids (*Balanus glandula* Darwin, 1854 and *Balanus crenatus* Bruguière, 1789), using their stylet to inject some paralyzing venom, then eating the soft tissues and leaving the empty carapace behind. That study, however, only described feeding occurrences in wild-caught planktonic larvae, and did not follow the development. We address this gap in the knowledge of the life history of these ubiquitous marine predators, by raising newly hatched larvae of *Emplectonema viride* in the laboratory and documenting their feeding behavior, development, pelagic duration, and metamorphosis.

Material and Methods

Sampling of ripe adults

Clusters of acorn barnacles were collected from three different locations on the Oregon Coast during the months of October, November, January, February and May of 2019/2020 (Table 1). The clusters were taken to the laboratory, placed in glass containers and soaked in filtered seawater until worms crawled out. Worms were removed and kept in 150ml glass dishes in a sea table with running seawater at ambient sea temperature (12–15 °C). The worms were then observed under an Olympus stereoscope to assess the presence of mature gonads. Immature specimens were released at the same collection site.

Table 1. Sampling locations and time of collection of ripe *Emplectonema viride* along the Oregon Coast

Sampling location	GPS coordinates	Month of sampling
Charleston Marina	43° 20.63'N 124° 19.38' W	November, February, May
OIMB Boathouse dock	43° 20.96'N 124° 19.80'W	October, May
Bastendorff Beach	43° 21.10'N 124° 20.65'W	November, February

About 20 mature animals were kept in the laboratory in 150ml glass bowls with frequent water change (2–3 times a week) and placed in a seatable to maintain the temperature. Since there is no known cue to induce egg release in hoplonemerteans, the bowls were checked for released eggs once a day. Once a spawning female was spotted, the eggs were collected with a glass pipette, washed twice in filtered seawater to remove any contaminants and placed in a clean bowl with filtered seawater. Eggs were fertilized *in vitro* by dissecting ripe males, and adding a dilute suspension of sperm to eggs. The fertilized eggs in small glass dishes were placed on a thermoelectric cold plate at 12ºC for the next 24 hours. During this period, at every

hour, some embryos were gently compressed between a glass slide and a cover slip and photographed using a Spot 5.2 camera mounted on an Olympus BX51, equipped with DIC optics. After 24 hours the developing embryos were transferred to a 150ml glass bowl and placed in a sea table at ambient sea temperature (12–15 °C) and observed regularly (every 1 – 2 days) to check the stage of development.

Once the stylet developed, putative prey items were added to the cultures and the behavior of about 400 larvae (94 in our first attempt, and about 300 in posterior cultures) was observed under a microscope stereoscope. We chose common crustaceans present in the plankton (calanoid and cyclopoid adults and nauplii copepods; balanidae and pollicipedidae nauplii; and decapod zoea) to test feeding preference and observe feeding mechanism. The amount of prey added was adjusted to the consumption by *E. viride* larvae to maintain food availability (about 10 nauplii per *E. viride* larva). Once the larvae reached metamorphosis, newly settled barnacles were also offered as food. The barnacles were obtained live in groups on small rock chips chiseled from nearby shores. Some of the feeding events were recorded using a Point Grey Grasshopper 3 camera operated by StreamPix 7, mounted on an Olympus BX51, equipped with DIC microscope for smaller, younger, larvae or on a Leica Z6 Apo macroscope for bigger, older larvae and juveniles.

Two to three larvae were observed and photographed using a Spot 5.2 camera mounted on an Olympus BX51, equipped with DIC optics once a day for the first 10 days of development, and then twice a week until the first larva of the culture reached juvenile stage. After metamorphosis, the cultures were checked once a week.

Confocal microscopy

To document developmental progress, 12 larvae were preserved for confocal microscopy right after hatching (2 days after fertilization) and successively older stages when modifications usually happen (4, 6, 8, 30 and 60 days, and juvenile). We only fixed larvae after hatching due to methodological constrains. The larvae were stained using the fluorescent marker phalloidin, which binds to actin filaments, and examined using confocal microscopy (Maslakova and von Dohren 2009). Briefly, about three stained larvae were placed in a slide coated with 1% poly-L-lysine and immersed in either 1X PBS or 90% glycerol, covered with a cover slip and sealed with nail polish. To better visualize internal structures three larvae of each stage were clarified with Murray clear and mounted as the others.

All slides were observed using an Olympus Fluoview 1000 confocal system mounted on an Olympus IX81 inverted microscope and imaged with 20X 0.85 NA, 40X 1.3 NA, and 60X 1.4 NA oil-immersion lenses. Each larva was scanned in 0.5μm focal steps and the confocal stacks were further processed using ImageJ (Wayne Rasband, National Institutes of Health, Bethesda, MD).

Results

Larval development

Ripe females of *Emplectonema viride* have a pinkish to brownish color in the ventral surface due to the many oocytes present in each ovary. Females tended to spawn right after water change, but sometimes they would spawn with no apparent cue. Spawned oocytes are round, opaque, and pinkish. Oocytes of one spawning event measured between 110 and 140 μm in diameter (n=9), and were surrounded by a chorion (131–160 μm) and jelly coat (204–351 μm). Fertilized eggs completed meiosis and then underwent equal spiral cleavage (Fig. 1A–G). At 12–14 ºC, the first polar body

appeared 25 minutes after fertilization; first cleavage was observed at 90 minutes, second cleavage after 140 minutes, third cleavage after 200 minutes and fourth cleavage after 240. Olive-shaped, uniformly ciliated, larvae with a thin apical tuft hatched after 38 hours. Two-day-old larvae swam actively, had a prominent apical tuft and a thin caudal ciliary cirrus (Fig. 2A). The rudiments of the proboscis, dorsal commissure and lateral nerve cords could be identified by confocal microscopy in two-day-old larvae (Fig. 3A). By the fourth day functional musculature, midgut and foregut were already developed (Fig. 3B), as were the first two ocelli and stylet (Figs. 2B; 4A), when they start feeding. When relaxed the larvae were about 200 μm long and 110 μm wide.
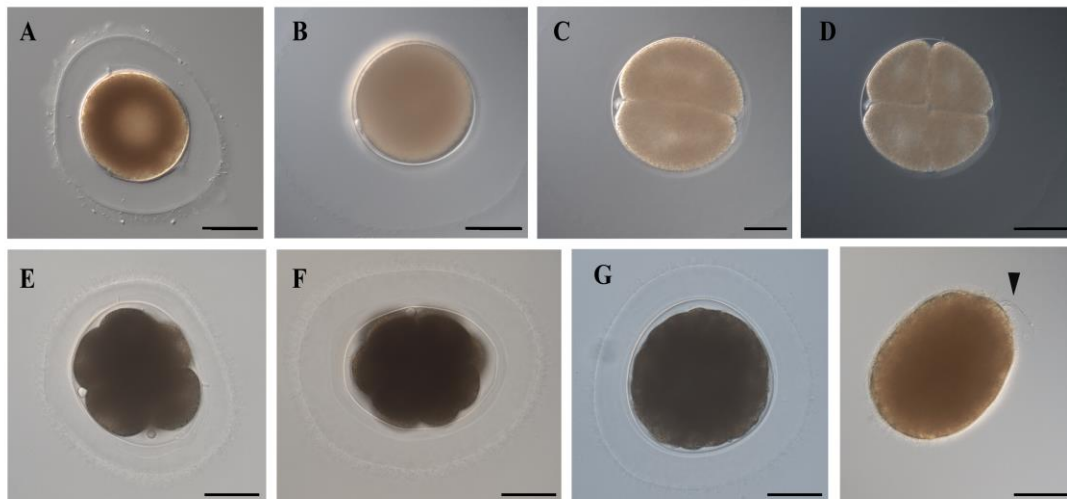
Figure 1. Spawned egg with tight chorion of *Emplectonema viride*. A: unfertilized egg; B: polar body formation; C: 2-cell stage; D: 4-cell stage; E: 8-cell stage; F: 16-cell stage; G: 32-cell stage; H: 36-hours post fertilization hatchling. Arrow indicates apical tuft. Scale bars: 50μm.

Figure 2. Initial development of *Emplectonema viride* larva. A: 2-day post fertilization, arrow indicates the apical tuft; B: 4-day post fertilization, arrow indicates the stylet; C: 6-day post fertilization; D: 12-day post fertilization, arrows indicate the opening of cerebral organs; E: 15-day post fertilization with four ocelli; F: 18-day post fertilization, with six ocelli; G: 31-day post fertilization, note the change in color. Scales bar: 50 μm (A, B); 100 μm (C–F); 200 μm (G).

About six days after fertilization, the larvae assumed a more elongated shape and a darker color in the gut (Fig. 2C). By the eighth day confocal microscopy revealed the accumulation of oil droplets in the posterior region. We did not detect the invaginations that originate the cerebral organs, but the cerebral organ openings appeared around day 12 (Fig. 2D). The second pair of ocelli appeared around day 15 and the third around day 18 (Fig. 2E–F). Around day 15 the larvae also acquired a greenish color in the epidermis, becoming darker as they grew (Fig. 2F–G). Between day 15 and day 30, no changes were perceived in the larvae, except for increase in size. 30-day-old larvae had a more developed gut, very dark epidermis and an elongated shape similar to the ones found in plankton tows (Fig. 5A). The number of ocelli, however, did not change. At 45 days the larvae had four pairs of ocelli, but

sometimes the ocelli were not paired, and the morphology of the stylet already showed similarities with the adult stylet, but with different proportions (Fig. 4B). The fifth pair of ocelli appeared around the 76th day after fertilization (Fig. 5B), when the midgut started to show invaginations that would become the midgut diverticulum (Fig. 5F). At 76 days after fertilization, the larvae presented six pairs of ocelli and many invaginations at the midgut (Fig. 5C, G).
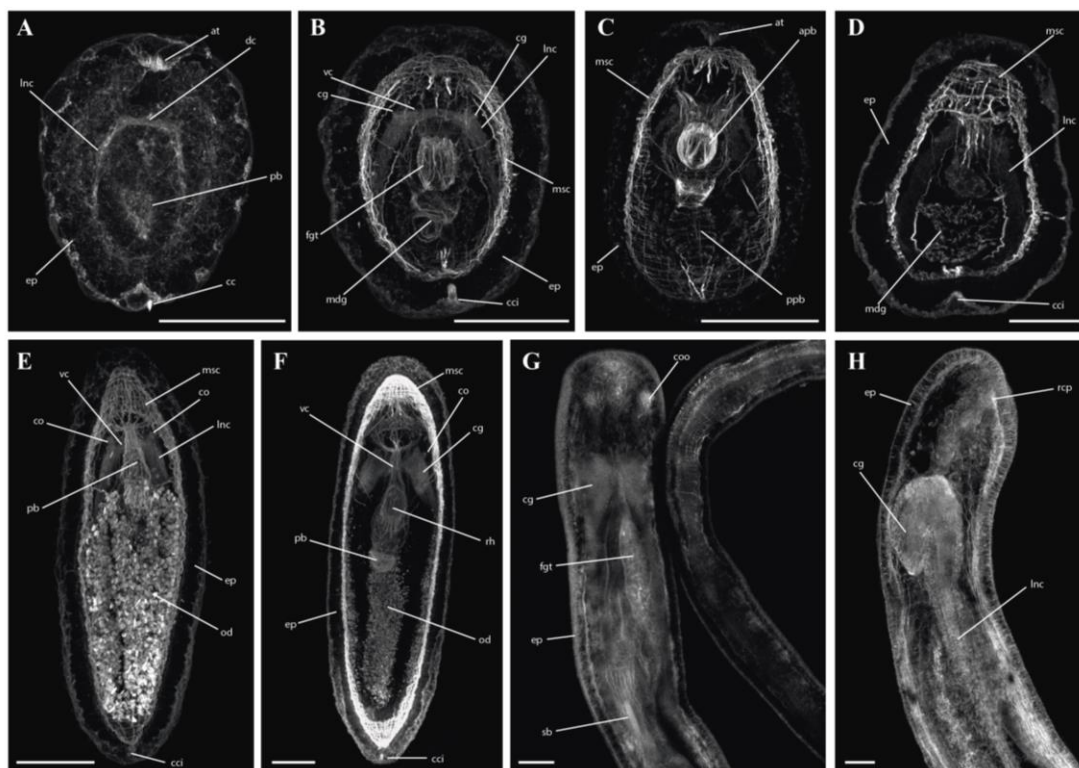
Figure 3. Confocal Z projections of substacks chosen to illustrate major morphological structures of phallacidin-labeled planktonic larva and juvenile of *Emplectonema viride*. A: 2-day post fertilization; B: 4-day post fertilization, dorsal view; C: 4-day post fertilization; D: 6-day post fertilization; E: 8-day post fertilization; F: 30-day post fertilization; G: juvenile, ventral view; H: juvenile, lateral view. apb: anterior proboscis; at: apical tuft; cc: caudal cirrus; cci: caudal cirrus insertion; co: cerebral organs; coo: cerebral organs openings; dc: dorsal commissure; ep: epidermis; fgt: foregut; lnc: lateral nerve cord; mdg: midgut; msc: musculature; od: oil droplets; pb: proboscis; ppb: posterior proboscis; rcp: rhynchostomopore; rh: rhynchocoel; sb: stylet basis; vc: ventral commissure. Scale bars: 50 μm.
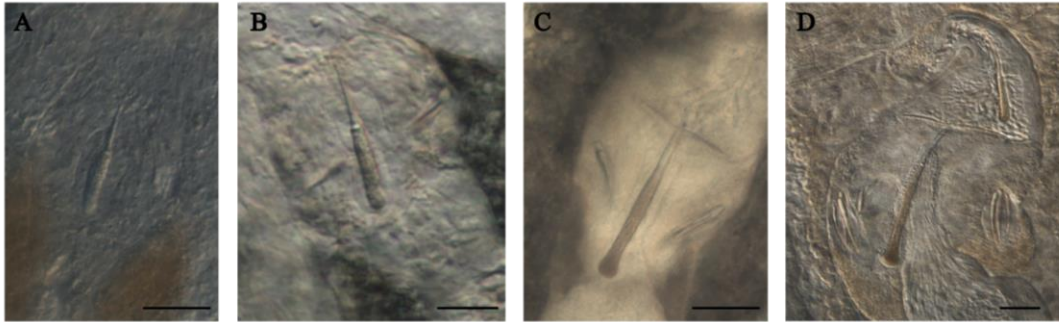
Figure 4. Stylets of *Emplectonema viride* in different stages of development. A: 4-day post fertilization central stylet and basis; B: 25-day post fertilization central and accessory stylets; C: 45-day post fertilization central and accessory stylets and basis; D: juvenile central and accessory stylets and basis. Scale bars: 20 μm (A–B); 50 μm (C–D).

Figure 5. Development of *Emplectonema viride* planktotrophic larva. A: general view of 43-day post fertilization larva; B: anterior region of 76-day post fertilization larva; C: anterior region of 100-day post fertilization larva; D: anterior region of 109-day post fertilization juvenile; E: general view of 109-day post fertilization juvenile; F: posterior region of 76-day post fertilization larva; G: posterior region of 100-day post fertilization larva; H: posterior region of 109-day post fertilization juvenile; I: larva posterior region; J: juvenile posterior region for comparison (Note the difference in color between the different stages and the lack of caudal cirrus and long cilia in the juvenile). Scale bars: 200 μm (A,B, D, F, I, J); 100 μm (C, G,H); 500 μm (E).

The first juvenile appeared at 109 days after fertilization, and the average time for metamorphosis was about 120 days. We considered a worm to be a juvenile when

they no longer exhibited positive phototaxis and started to crawl the majority of time. The juvenile had a very pale epidermis, but still six pairs of ocelli (Fig. 5D–E). The animals at this stage already presented many diverticula along the entire midgut (Fig. 5H). Subtle morphological changes accompany the behavioral changes that we refer to as settlement: juveniles lose the caudal cirrus, which is present in all larval stages, and the body ciliation shortens dramatically (Fig. 5I–J). The basis and stylet in juvenile worms also assumed similar proportions as the adults (Fig. 4C).
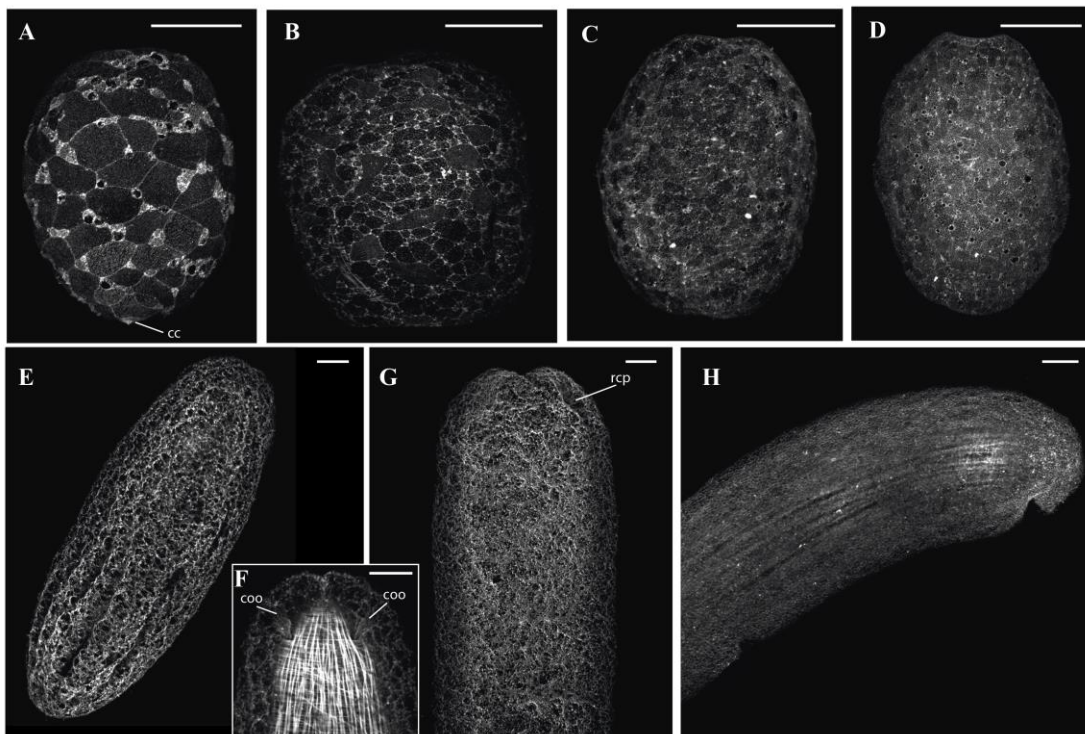
Figure 6. Confocal Z-projections of *Emplectonema viride* larva in different stages of development and juvenile showing the outlines of epidermal cells and gradual replacement of the transitory larval epidermis (large cells) by the definitive epidermis. A: 2-days post fertilization; B: 4-days post fertilization; C: 6-days post fertilization; D: 8-days post fertilization; E: 30-days post fertilization; F: 30-days post fertilization, focusing on the cerebral organ openings; G: 60-days post fertilization; H: juvenile. cc: caudal cirrus; coo: paired cerebral organs openings; rcp: rhynchostomopore. Scale bars: 50 μm.

Early development from hatchling to feeding larva was accompanied by dramatic changes in the epidermis, similar to that described in the direct-developing larva of

Paranemertes peregrina. The epidermis in the 2-day old *E. viride* larva was composed of a discrete number of large cells with groups of a few small cells in interstices between the large ones (Fig. 6A). The large cells were progressively substituted as the larva grew (Fig. 6A–F). In the 4-day old larva the large cells remain but the epidermis is dominated by increasing numbers of small cells (Fig. 6B); the large cells almost disappeared in the 8-day old larva (Fig. 6D). In the 30-day old larvae, the epidermis is composed mainly of glandular and ciliated cells (Fig. 6E–F). The epidermis of the juvenile, however, is smother, consisting of very small cells (Fig. 6G).

Feeding behavior and prey preference

The larvae started to feed as soon as they developed a stylet, at 4-days old, and kept on feeding until metamorphosis. They showed a near-absolute preference for barnacle nauplii and cyprids; we observed only one feeding event upon a calanoid copepod. This copepod was not identified because there was not tissue left over after the feeding event. The nauplii and cyprids, however, were identified as *Balanus glandula* and *Balanus crenatus*, the two most common species of acorn barnacle in the region.

Larvae reacted almost instantly to the addition of prey in the culture. They started to swim more actively and change their body shape to something resembling a tadpole (Fig. 7A; supplemental video 1). However, they were not observed to attack barnacle larvae in this shape. Most observed encounters occurred when the *E. viride* larva swam slowly near the bottom and attacked barnacle larvae already trapped by mucus laid by the *E. viride* larvae. In some cases, they seemed to sense prey caught by a line of trailing mucus, curving the head towards the tail, coiling around the prey, and then attacking using the stylet (Supplemental video 2).
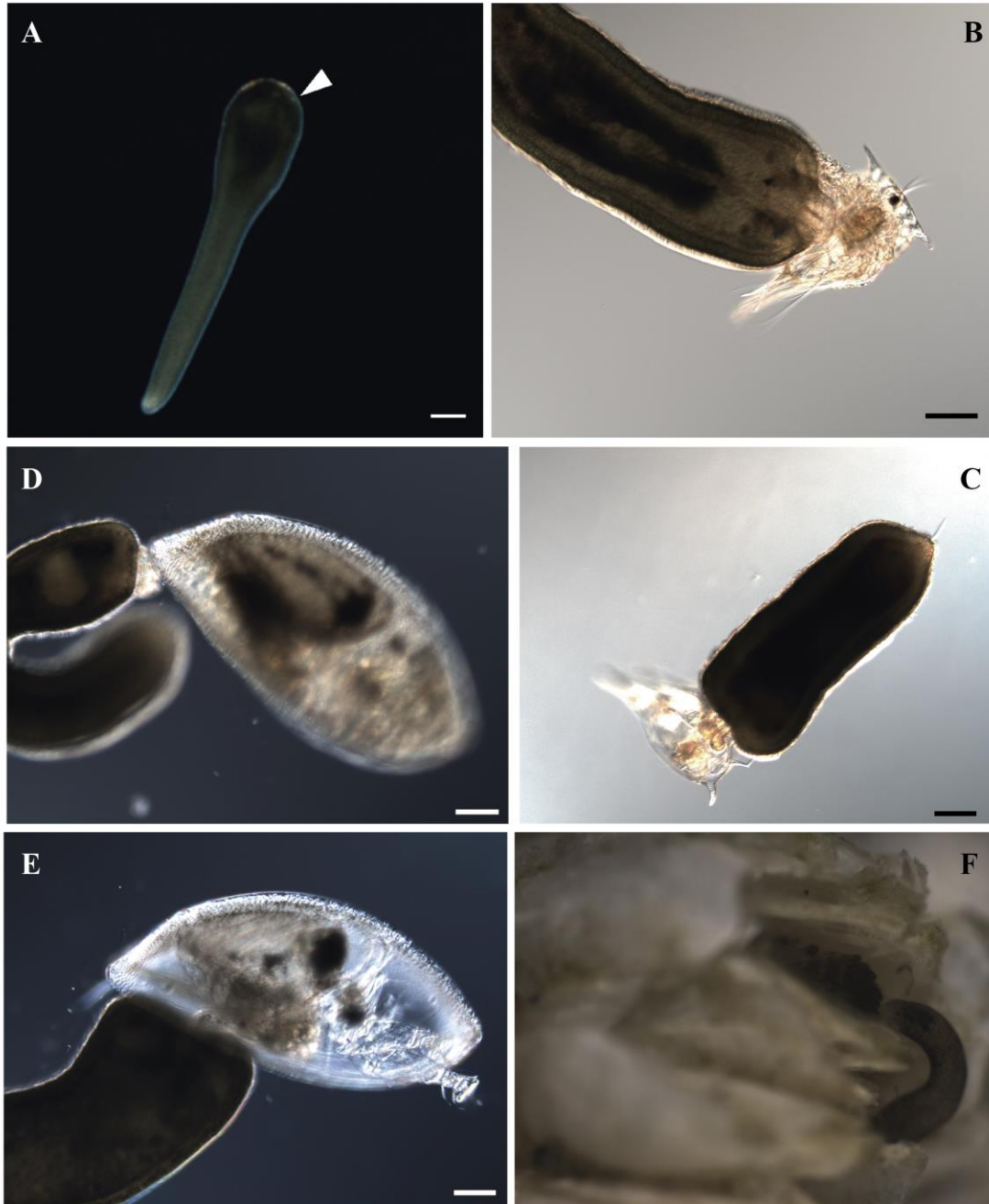
Figure 7. Feeding behaviour of *Emplectonema viride* larvae (A-E) and juvenile (F). A: planktotrophic larva swimming in tadpole shape. Arrow indicates the paired cerebral organ openings; B: larva attacking a barnacle nauplius; C: larva feeding upon a barnacle nauplius; D: larva attacking a barnacle cyprid; E: larva feeding upon a barnacle cyprid; F: juvenile feeding upon a newly settle *Balanus glandula* in 3.5 zoom. Scale bars: 50 µm (A); 100 µm (B–E).

The whole predation event took between 5 to 15 minutes, taking sometimes longer for

smaller *E. viride* or when they consumed larger cyprids. As stated above, the attacks

on nauplii usually initiated after the nauplii was already caught by the mucus produced by *E. viride* larvae. Usually, in a successful attack, *E. viride* would first attack the anterior region of a nauplius, around the junction of the first antennae and the carapace, using the proboscis and perforating the prey many times (Fig. 7B; supplemental video 3). Following attack, the prey slowly ceased moving, while the larvae was still coiled around or, when smaller, swimming around the prey. Once prey movements reduced to mere spasms, *E. viride* larva inserted the proboscis inside the carapace and sucked in all tissues, including the naupliar eye, typically leaving behind only naupliar gut contents (Fig. 7C, supplemental video 4). If the *E. viride* larva was disturbed during this process, it would abandon the prey. Therefore, most recorded encounters happened when both *E. viride* larvae and the nauplii were already between a slide and cover slip. As a result, the recorded feeding events were less successful and took longer. The larvae, however, had no objections to sharing the same prey item with other individuals. When smaller, many *E. viride* could attack the same nauplius.

All the observed feeding events upon cyprids (n = 12 events) had exactly the same sequence. First, the cyprid was caught in the mucus already present in the culture, which restrained some of its movements, and slows it a little. After this, the *E. viride* larva coiled around the cyprid and attacked between the valves, in the posterior region, inserting the proboscis and stylet many times (Fig. 7D). After being attacked, the cyprid moved antennules and sometimes thoracic appendages, but did not jump away or make any other greater movement. The *E. viride* larva then prowled around the cyprid and, after about 1–2 minutes, inserted the proboscis at the same region as before and began to suck in all tissues, including the eye and all oil droplets (Fig. 7E;

supplemental video 5 [at 2x normal speed]). The ingestion would take about 5–8 minutes.

Juvenile development and first reproductive adults

In October 2019 two cultures were successfully started, with 94 larvae in total (a first spawn with 23 embryos and a second one with 71). From these, 31 worms reached metamorphosis, while 56 specimens were fixed during their development. Once individuals reached juvenile state, we offered recently-settled (hence small) barnacles to test for a change in diet. At first no feeding was observed, but the juveniles hid under rock chips bearing barnacles. Some empty barnacle shells were spotted after one day or two, but no feeding was witnessed until the cultures were observed under dim light. The feeding events observed took place after several minutes (up to 120 minutes) of monitoring. The worms crawled around the barnacle mats and attacked only barnacles that were alive but not moving vigorously. The worms initiated the attacks by inserting the proboscis between the operculum and the marginal plates or through the apertures, and apparently pierced the tissues with the stylet many times. Some seconds after the injection, the barnacle stopped moving and the worm inserted the proboscis through the operculum opening and started suck out barnacle tissues (Fig. 7F; Supplemental video 7), leaving behind the empty capitulum after five to ten minutes. During feeding, it was possible to see liquefied barnacle tissues entering the worm's gut.

From the 31 worms that reached juvenile state, 26 kept on growing under barnacle diet (five died after trying to escape the bowls). These 26 juveniles grew at different rates (varying around 2.5x in the same larval culture), the epidermis becoming darker as they grew. About eight months after hatching, these young adults started to show

40

maturing gonads, and after about 45 more days they spawned. The spawned eggs had the same characteristics as the ones spawned by their wild counterparts: round, opaque, and pinkish, and with similar size range.

Discussion

This is the first time a hoplonemertean has been raised in laboratory conditions from fertilized egg to reproductive adult. The initial development of *E. viride* follows the general pattern described in other hoplonemerteans, except for invaginations that originate the cerebral organs, such as *Paranemertes peregrina* Coe, 1901 and *Quasitetrastemma stimpsoni* Chernyshev, 1992 (Maslakova and von Doehren, 2009) and *Pantinonemertes californiensis* (Hiebert et al., 2010), in that a transitory epidermis of large multiciliated cells covers the hatchling, to be progressively replaced by a definitive of smaller cells. *Emplectonema viride* embryos, however, have a more accelerated development considering they develop all feeding apparatus, including central stylet, within four days after fertilization, while truly lecithotrophic species only develop such apparatus up to three weeks after fertilization (Stricker 1985; Chernyshev 2008; Maslakova and von Dohren 2009). However, the absence of a stylet is not always an evidence of true lecithotroph, as seen in *Carcinonemertes errans* larvae (von Dassow *et al.* in press). This condensed early development might be why we were not able to observe the invaginations that become the cerebral organs. These invaginations occurred around the second day after fertilization in *P. peregrina* (Maslakova & von Doren, 2009). However, they were already not visible in *E. viride* hatchlings, 36 hours after fertilization. In this stage, the internal anatomy was already similar to the 4-day old larvae of *P. peregrina*.

It is noteworthy that *E. viride*'s larvae feed on the larvae of the same animals – barnacles – that *E. viride* preys upon as adults. It could be argued that this is a mere consequence of abundance and susceptibility, as a) barnacle nauplii are dominant members of the local plankton and don't swim as fast as other abundant crustaceans such as copepods, and b) barnacle adults are among the most abundant, yet the least mobile, of benthic crustaceans. Alternatively, it could reflect a true selectivity. Although we offered a variety of prey items, the amount of each prey item was still reflected in this microcosm. Some hoplonemerteans species are known to have specific feeding preferences, however when their preferred item is not available they can prey on alternative organisms (Roe 1993; Thiel and Kruse 2001).

The initial development of *E. gracile* is well described in Iwata (1960) and is very similar to what was observed for *E. viride* in this study. Although *E. gracile* embryos hatched earlier and no feeding was observed, the larvae already have a fully developed feeding system within eight days, dying about 17 days after hatching. This is consistent with a possible planktotrophic stage, a possibility also raised by Iwata (1960), and by Chernishev (2008) when studying the development of *Quasitetrastemma* species.

Consequences of a feeding larva for hoplonemertean dispersion

Few hoplonemertean species have been observed until metamorphosis, while most of those observed died within a few days after hatching (e.g. Iwata 1960; Stricker and Reed 1981; Chernyshev 2008). Once we know many juvenile-like hoplonemertean hatchlings are actually carnivorous larvae, it seems highly probable that these species, as well as many other hoplonemertean species, also have a feeding planktonic stage before settling, which is also suggested by other researchers, like Iwata (1960).

Previous molecular and observational works raised the hypothesis of planktotrophy to explain the results found (e.g. Andrade *et al.* 2011; Tulchinsky *et al.* 2012; Maslakova & Hiebert, 2014; Mendes *et al.* 2018). By raising *E. viride* on a carnivorous diet in the lab, we are able to assert that they feed and grow solely on such means, confirming such hypothesis.

Oregon populations of *E. viride* harbor low genetic diversity, with only two haplotypes present in this area (Mendes *et al.* 2021). The long larval period, which lab culture suggest may be as long as 120 days has a direct influence on gene flow within and between populations, since the adults do not have high dispersal potential. The adult worms live among incrusting communities, swimming actively only occasionally. The importance of larval dispersal for population connectivity in benthic species is recognized in many population genetic studies, and larval planktonic duration is directly linked to the amount of gene flow between populations (Palumbi 1994; Bohonak 1999; Hellberg *et al.* 2002; Hellberg 2007; Kelly and Palumbi 2010; Selkoe and Toonen 2011). Selkoe and Toonen (2011), gathering data from many previous papers, were able to show graphically how the duration of the larval phase has direct impact on the distance achieved by the larvae (inferred from the pelagic larval duration), and consequently on the population structuring of these species. Therefore, the Pacific populations of *E. viride* are probably very well connected, as seen in the Oregon populations. The sister species, *E. gracile*, also shows well connected haplotypes throughout the Northeast Atlantic Ocean until the North Sea (Mendes *et al.* 2021), which suggests a planktotrophic larva in this species as well.

A long planktotrophic larval phase seems to be more widespread among hoplonemertean species than previously thought, indicating that the intersection of life history evolution and biodiversity among nemerteans deserves a closer look.

Acknowledgments

References

Andrade, S. C. S., Norenburg, J. L., & Solferini, V. N. (2011). Worms without borders: genetic diversity patterns in four Brazilian *Ototyphlonemertes* species (Nemertea, Hoplonemertea). *Marine biology*, *158*(9), 2109.

Andrade, S.C., Montenegro, H., Strand, M., Schwartz, M.L., Kajihara, H., Norenburg, J.L., Turbeville, J.M., Sundberg, P. & Giribet, G. (2014). A transcriptomic approach to ribbon worm systematics (Nemertea): resolving the Pilidiophora problem. *Molecular biology and evolution*, *31*(12), pp.3206-3215.

Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *The Quarterly review of biology*, *74*(1), 21-45.

Bowen, B. W., Bass, A. L., Muss, A., Carlin, J., & Robertson, D. R. (2006). Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): exploring

links between pelagic larval duration and population connectivity. *Marine Biology*, *149*(4), 899-913.

Chernyshev, A. V. (2008). Larval development of nemerteans of the genus *Quasitetrastemma* (Nemertea: Monostilifera). *Russian Journal of Marine Biology*, *34*(4), 258-262.

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual review of marine science*, *1*, 443-466.

Gibson, R. (1995). Nemertean genera and species of the world: an annotated checklist of original names and description citations, synonyms, current taxonomic status, habitats and recorded zoogeographic distribution. *Journal of Natural History*, *29*(2), 271-561.

Griffin, B. B. (1898). Description of some marine nemerteans of Puget Sound and Alaska. *Annals of the New York Academy of Sciences*, *11*(1), 193-218.

Hellberg, M. E., Burton, R. S., Neigel, J. E., & Palumbi, S. R. (2002). Genetic assessment of connectivity among marine populations. *Bulletin of marine science*, *70*(1), 273-290.

Hellberg, M. E. (2007). Footprints on water: the genetic wake of dispersal among reefs. *Coral Reefs*, *26*(3), 463-473.

Hiebert, L. S., Gavelis, G., von Dassow, G., & Maslakova, S. A. (2010). Five invaginations and shedding of the larval epidermis during development of the hoplonemertean *Pantinonemertes californiensis* (Nemertea: Hoplonemertea). *Journal of Natural History*, *44*(37-40), 2331-2347.

Hiebert, T. C., Von Dassow, G., Hiebert, L. S., & Maslakova, S. A. (2013, April). Long-standing larval mystery solved *Pilidium recurvatum* is the larva of *Riserius* sp., a basal heteronemertean (Heteronemertea; Pilidiophora; Nemertea). *Integrative and Comparative Biology*, *53*(1), E92.

Hiebert TC, Maslakova S (2015) Integrative taxonomy of the *Micrura alaskensis* Coe, 1901 species complex (Nemertea: Heteronemertea), with descriptions

of a new genus *Maculaura* gen. nov. and four new species from the NE Pacific. Zoological science, 32(6): 615-637.

Hunt, M. K., & Maslakova, S. A. (2017). Development of a lecithotrophic pilidium larva illustrates convergent evolution of trochophore-like morphology. *Frontiers in Zoology*, *14*(1), 1-18.

Iwata, F. (1960). Studies on the comparative embryology of nemerteans with special reference to their interrelationships. *Publications from the Akkeshi Marine Biological Station*, *10*, 2-51.

Jägersten, G. (1972). Evolution of the Metazoan Life Cycle First Printing edition.

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

Laumer, C. E., Fernández, R., Lemer, S., Combosch, D., Kocot, K. M., Riesgo, A., ... & Giribet, G. (2019). Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proceedings of the royal society B*, *286*(1906), 20190831.

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.

Levin, L. A., Caswell, H., DePatra, K. D., & Creed, E. L. (1987). Demographic consequences of larval development mode: planktotrophy vs. lecithotrophy in Streblospio benedicti. *Ecology*, *68*(6), 1877-1886.

Maslakova, S. A., Martindale, M. Q., & Norenburg, J. L. (2004). Fundamental properties of the spiralian developmental program are displayed by the basal nemertean *Carinoma tremaphoros* (Palaeonemertea, Nemertea). *Developmental Biology*, *267*(2), 342-360.

Maslakova, S. A., & Malakhov, V. V. (1999). A hidden larva in nemerteans of the order Hoplonemertini. In doklady biological sciences section c/c of doklady-akademiia nauk sssR (Vol. 366, pp. 314-317). Nauka/interperiodica publishing.

Maslakova, S. A., & von Dohren, J. (2009). Larval development with transitory epidermis in *Paranemertes peregrina* and other hoplonemerteans. *The Biological Bulletin*, *216*(3), 273-292.

Maslakova, S. A. (2010). Development to metamorphosis of the nemertean pilidium larva. *Frontiers in Zoology*, *7*(1), 30.

Mendes, C. B., Norenburg, J. L., Solferini, V. N., & Andrade, S. C. S. (2018). Hidden diversity: Phylogeography of genus *Ototyphlonemertes* Diesing, 1863 (Ototyphlonemertidae: Hoplonemertea) reveals cryptic species and high diversity in Chilean populations. PloS one, 13(4), e0195833.

Mendes, C. B., Delaney, P., Turbeville, J. M., Hiebert, T., & Maslakova, S. (2021). Redescription of *Emplectonema viride*–a ubiquitous intertidal hoplonemertean found along the West Coast of North America. *ZooKeys*, *1031*, 1.

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

Podsiadlowski, L., Braband, A., Struck, T. H., von Döhren, J., & Bartolomaeus, T. (2009). Phylogeny and mitochondrial gene order variation in Lophotrochozoa in the light of new mitogenomic data from Nemertea. *BMC genomics*, *10*(1), 364.

Roe, P., 1993. Aspects of the biology of *Pantinonemertes californiensis*, a high intertidal nemertean. *Hydrobiologia 266(1)*, 29–44.

Thollesson, M., & Norenburg, J. L. (2003). Ribbon worm relationships: a phylogeny of the phylum Nemertea. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1513), 407-415.

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.

Strand, M., Norenburg, J., Alfaya, J. E., Ángel Fernández-Álvarez, F., Andersson, H. S., Andrade, S. C., ... & Chernyshev, A. (2019). Nemertean taxonomy—Implementing changes in the higher ranks, dismissing Anopla and Enopla. *Zoologica Scripta*, *48*(1), 118-119.

Stricker, S. A., & Reed, C. G. (1981). Larval morphology of the nemertean *Carcinonemertes epialti* (Nemertea: Hoplonemertea). *Journal of Morphology*, 169(1), 61-70.

Stricker, S. A. (1985). The stylet apparatus of monostiliferous hoplonemerteans. *American Zoologist*, *25*(1), 87-97.

Stricker, S. A., & Norenburg, J. L. (2002). Phylum Nemertea. *Atlas of marine invertebrate larvae*, 163 - 177.

Struck, T. H., & Fisse, F. (2008). Phylogenetic position of Nemertea derived from phylogenomic data. *Molecular Biology and Evolution*, *25*(4), 728-736.

Thiel, M., & Kruse, I. (2001). Status of the Nemertea as predators in marine ecosystems. *Hydrobiologia*, *456*(1), 21-32.

Tulchinsky, A. Y., Norenburg, J. L., & Turbeville, J. M. (2012). Phylogeography of the marine interstitial nemertean *Ototyphlonemertes parmula* (Nemertea, Hoplonemertea) reveals cryptic diversity and high dispersal potential. *Marine Biology*, 159(3), 661-674.

von Dassow, G., Emlet, R. B., & Maslakova, S. A. (2013). How the pilidium larva feeds. *Frontiers in Zoology*, *10*(1), 47.

von Dassow, G., Mendes, C. B., Robbins, K., Andrade, S. C., & Maslakova, S. A. (2021). Hoplonemertean larvae are planktonic predators that capture and devour active animal prey. *bioRxiv* doi: 10.1101/2021.02.02.429399v1

Supplemental files

Supplemental video 1: Planktotrophic larva of *Emplectonema viride* swimming in tadpole shape at normal speed. Recorded at 2x zoom.

Supplemental video 2: Planktotrophic larva of *Emplectonema viride* attacking and feeding upon barnacle nauplius at 0.5x normal speed. Recorded at 4x zoom.

Supplemental video 3: Planktotrophic larvae of *Emplectonema viride* feeding upon the same barnacle nauplius at normal speed. Recorded at 20x zoom.

Supplemental video 4: Planktotrophic larva of *Emplectonema viride* attacking and piercing a barnacle nauplius with the stylet at normal speed. Recorded at 20x zoom.

Supplemental video 5: Planktotrophic larva of *Emplectonema viride* feeding upon a barnacle at 2x normal speed. Recorded at 10x zoom.

.

Chapter 2. Species delimitation integrative approach reveals three new species in the *Nemertopsis bivittata* complex

Cecili B. Mendes, Jon L. Norenburg & Sónia C. S. Andrade

Abstract.

The presence of cryptic species is fairly frequent in many invertebrate groups and even more so among invertebrates with simple morphology, such as nemerteans. Consequently, the use of molecular methods for species delimitation has become a needed tool to complement morphological analyses to better recognise such species. *Nemertopsis bivittata* is one example of species with subtle morphological variation, but ample geographic distribution, being a good candidate for a species complex study. Here we applied two mitochondrial genes, and 2903 SNP (Single Nucleotide Polymorphism) variants in addition to morphological characters to investigate the presence of cryptic species among specimens previously identified as *N. bivittata* along the Brazilian Coast. To do so, specimens were collected at 15 different sites in the north-east, south-east and southern regions. Three new species of *Nemertopsis* are described based on morphological and molecular analyses: *Nemertopsis caete* sp. nov., *Nemertopsis pamelaroeae* sp. nov. and *Nemertopsis berthalutzae* sp. nov. The species *N. pamelaroeae* and *N. berthalutzae* present broad distributions from north-east to south-east; *N. caete*, however, is restricted to the north-east coast. This is the first study to use this combined approach in nemerteans and shows the advantages of integrating genomic markers with classical taxonomy, and applying objective approaches to delimiting species as independently evolving entities.

Keywords: biodiversity, genetics, mitochondrial DNA, molecular taxonomy, Nemertea, South America, species delineation, taxonomy

Introduction

Knowledge about marine biodiversity has been rapidly expanding in the past decades, with the total number of valid species described at ~2.3 million and increasing every year (Lewin *et al*. 2018 and references therein). Recognition and description of new species is increasingly important to understanding the effects of climate change and oceanic acidification (Fowler 2015; Asch *et al*. 2018; Dee *et al*. 2019) on the distribution and extinction of organisms. However, due to difficult sampling access and decreasing taxon expertise, marine invertebrates encompass the greatest phyletic range of poorly studied animals, which includes understanding the nature and prevalence of cryptic species among those invertebrates (Sundberg *et al*. 2009*a*; Adams *et al*. 2014; Fišer *et al*. 2018). Molecular data have been of great assistance in discerning cryptic species by many approaches used in species delimitation studies (Sundberg *et al*. 2009*a*; Adams *et al*. 2014; Cornils *et al*. 2017; Bocek *et al*. 2019; Kobayashi and Sota 2019; Quattrini *et al*. 2019; Campbell *et al*. 2020). Among nemerteans, ~1500 described species of mainly marine worms (Strand *et al*. 2019), incorporating molecular data has become increasingly common and even indispensable for species description and sometimes for identification (Sundberg *et al*. 2009*b*, 2016*a*; Strand and Sundberg 2011; Junoy *et al*. 2011; Hiebert and Maslakova 2015; Strand *et al*. 2019, among others).

Nemertopsis bivittata (Delle Chiaje, 1841) is a widespread nemertean species described from the Mediterranean Sea, and recorded in both the Atlantic and Pacific Oceans (Corrêa 1955; Sánchez 1973; Gibson 1995). This is a free-living species that inhabits mussel and barnacle beds growing on natural and artificial substrates. It has a whitish body and two dorsal longitudinal dark lines. The species has a controversial taxonomic history, with many synonyms being described over the years (Gibson 1995; Norenburg *et al*. 2020). Recently, with the addition of molecular data for specimens from different parts of the world,

researchers have noticed that the morphological species actually comprise a species complex with many different species (Caplins *et al*. 2012; Turbeville, pers. comm.). We call this the *N. bivittata* species complex, which is distributed worldwide and comprises the many cryptic species identified as *N. bivittata* characterised by the dark longitudinal lines. Having that in mind, we collected putative *N. bivittata* specimens along the Brazilian coast and used an integrative approach to test for a species complex. The mitochondrial regions cytochrome *c* oxidase subunit I (*COI*) and *16S* rRNA (henceforth, *16S* rRNA), genomic (with 2903 SNPs) and morphological data were applied in a species delimitation analysis. That leads us to describe three new species of the *N. bivittata* complex from the Brazilian coast, using the concept of species as independently evolving entities.

Material and methods

Sampling

We collected samples of oyster beds in 15 localities, in six of the federated states along the Brazilian coast (Table 1, Fig. 1), during the years 2017–2019. Clusters of oyster bed were extracted from exposed rocks or pillars with a spatula, and transported to the laboratory in plastic bags. Once in the laboratory, the clusters were placed in plastic trays and covered with seawater for at least 2 h. After that period, oysters were separated and opened, and any worms found were carefully removed with a brush, placed in a dish, relaxed in a mixture 1:1 of $MgCl_2$ (stock approximately isotonic to seawater) and seawater, measured and photographed. Photographs of the head and body were taken with a stereoscope, and compressed specimens (head, stylet, and gonads when present) were photographed with a light microscope. Subsequently, most specimens were fixed in 99% ethanol; three specimens were fixed in 10% formalin and later stored in 70% ethanol. Specimens fixed in 99% ethanol were kept at $-20°C$ until DNA extraction. Holotypes and paratypes were deposited in the

invertebrate collection at the Museum of Zoology of the University of São Paulo (MZUSP).

Animals were collected under permits issued by Institute Chico Mendes (ICMBio), nos

55701 and 67004.

Table 1. Sampling locations, GPS coordinates and number of specimens found. Abbreviation for each location is shown in parentheses

Sampling location	State	Latitude	Longitude	Date of collection	Number of specimens
Paracuru (PNJ)	Ceará	–3.3997	–39.0137	Jan. 2017	2
Sabiaguaba (SBNS)	Ceará	–3.7687	–38.4358	Jul. 2017	4
Prado (PANS and PANJ)	Alagoas	–9.6769	–35.7515	Jan. 2019	38
Gamboa (GBN)	Espírito Santo	–20.8887	–40.7655	Jun. 2018	1
Itaoca (ItaNS and ItaNJ)	Espírito Santo	–20.9050	–40.7770	Jun. 2018	3
Itaipu (ITN)	Rio de Janeiro	–22.9741	–43.0471	Dec. 2017	2
Praia das Goiabas (MNS)	Rio de Janeiro	–23.0258	–44.5144	Feb. 2019 and May 2017	3
Prainha (PPN)	Rio de Janeiro	–23.1482	–44.6955	Feb. 2019	2
Barra do Corumbê (BCN)	Rio de Janeiro	–23.1791	–44.7112	Feb. 2019	2
Jabaquara (JNJ)	Rio de Janeiro	–23.2112	–44.7141	Apr. 2017 and Feb. 2019	39
Igararecê Marina (INJ and INES)	São Paulo	–23.7682	–45.4009	Apr. 2018	6
Pontal da Cruz Pier (PCN)	São Paulo	–23.7813	–45.3971	Feb. 2019	2
Araçá (ANC and ANE)	São Paulo	–23.8175	–45.4067	Jan. 2017 and Feb. 2019	15
Balneário dos Trabalhadores (PGN)	São Paulo	–23.8231	–45.4166	Feb. 2019	9
Ribeirão da Ilha (RNS)	Santa Catarina	–27.7223	–48.5642	May 2017	1
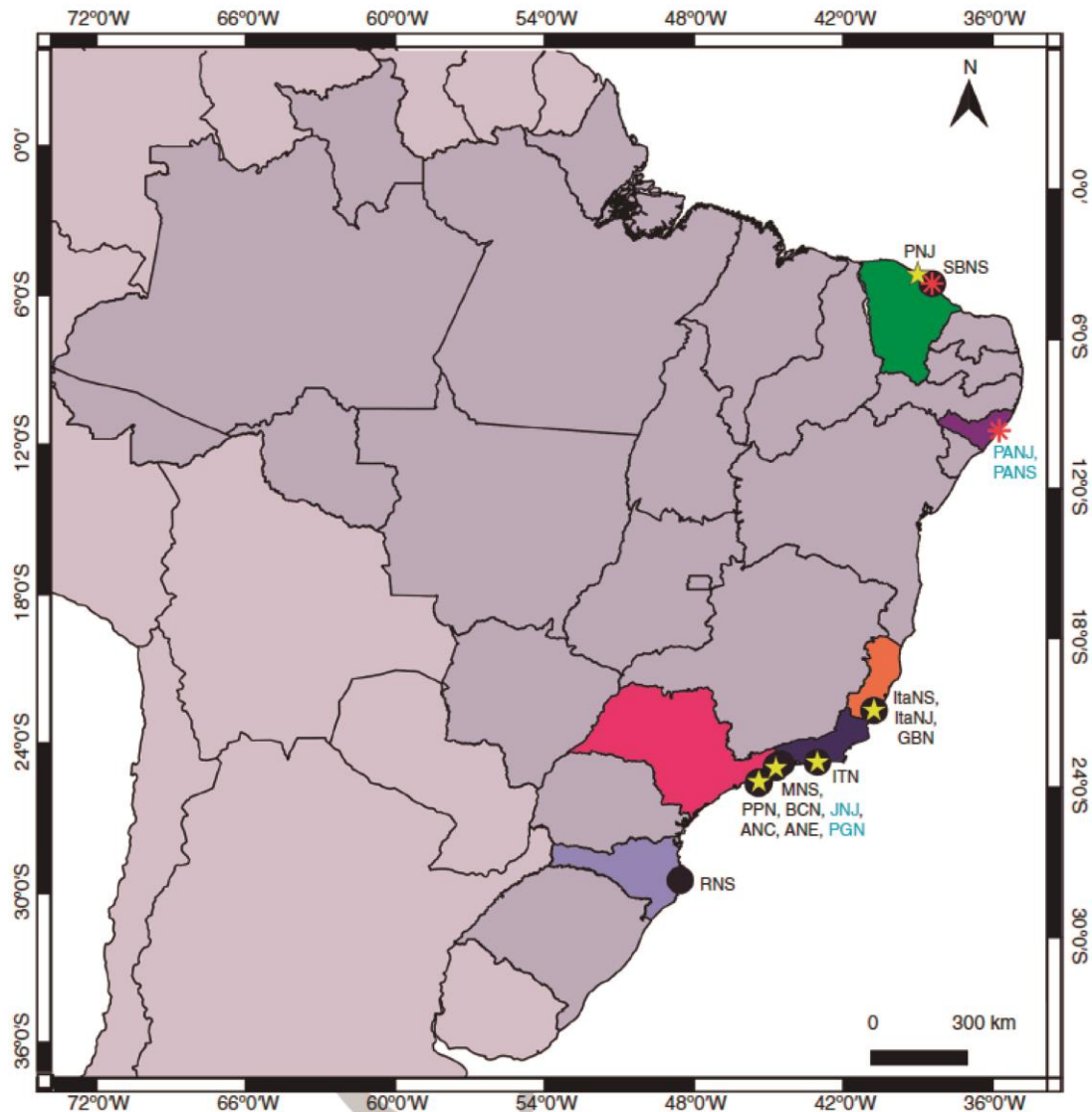
Figure 1. Sampling sites of *Nemertopsis* specimens along the Brazilian coast. Black circles: sampling sites of *Nemertopsis berthalutzae*, sp. nov. Red asterisks: sampling sites of *Nemertopsis caete*, sp. nov. Yellow stars: sampling sites of *Nemertopsis pamelaroeae*, sp. nov. Abbreviations in blue indicate the type locations of each species. Abbreviations for each sampling location: Paracuru (PNJ); Sabiaguaba (SBNS); Prado (PANS and PANJ); Gamboa (GBN); Itaoca (ItaNS and ItaNJ); Itaipu (ITN); Praia das Goiabas (MNS); Prainha (PPN); Barra do Corumbê (BCN); Jabaquara (JNJ); Igararecê Marina (INJ and INES); Pontal da Cruz Pier (PCN); Araçá (ANC and ANE); Balneário dos Trabalhadores (PGN); Ribeirão da Ilha (RNS). Each sampled state (Brazilian federative units) has a different colour: green (Ceará), dark purple (Alagoas), orange (Espírito Santo), blue (Rio de Janeiro), pink (São Paulo), light purple (Santa Catarina).

Morphological data

Specimens were observed alive under the stereoscope and the microscope, and character

data were recorded in accord with the character matrix of Sundberg *et al*. (2009*b*), using the

ecological and morphological characters: 1–3, 5–26, 40, 41, 45, 47, 54–60, 70, 104 and 105

from its matrix. In addition, pictures of 71 individual heads were selected for morphological

principal component analysis (PCA). A list of 71 pictures (one per specimen) was created

using tpsUtil (ver.1.74, https://tpsutil.software.informer.com/; Rohlf 2015) and tpsDig2 (ver.

2.31, https://tpsdig2.software.informer.com/; Rohlf 2015). This list was used as input for

MorphoJ (Klingenberg 2011), with six landmarks defined (Fig. 2*A*). The landmarks were

transformed using the option Procrusts fit (Dryden and Mardia 1998) to find optimal rotation

and better comparison between specimens. Transformed landmarks were used to create a

covariance matrix and this was used as input to PCA in MorphoJ.
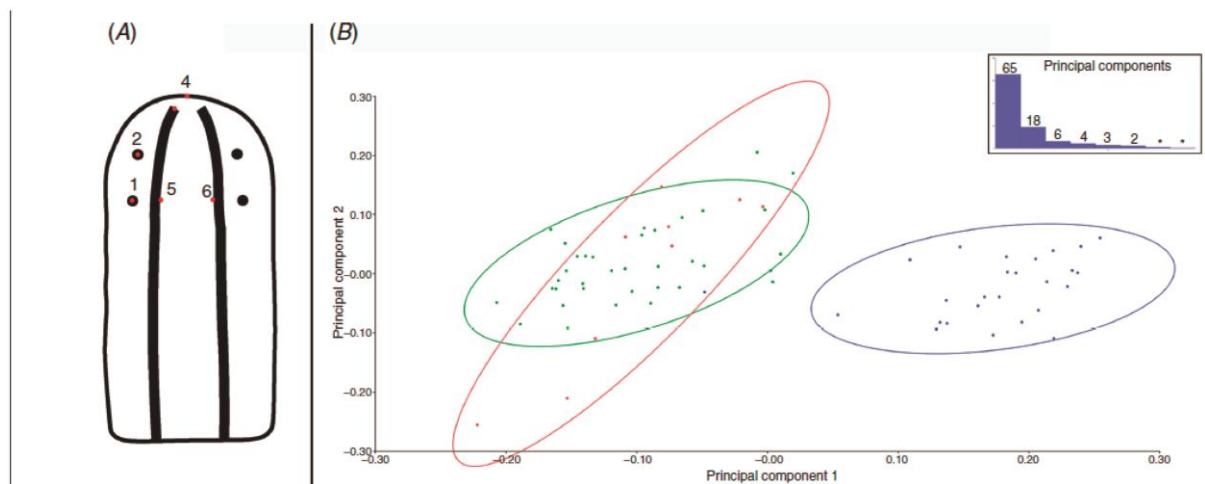
Figure 2. Landmarks used in the morphological principal component analysis (PCA) and PCA results. (*A*) Example of landmarks. 1, anterior pair of eyes, left side; 2, posterior pair of eyes, left side; 3, end of the dorsal lines, left side; 4, frontal margin; 5, inner side of left dorsal line at anterior pair of eyes level; 6, inner side of right dorsal line at anterior pair of eyes level. (*B*) PCA results from the first two PCs. Each dot represents one specimen. Blue dots, specimens of *Nemertopsis berthalutzae*, sp. nov.; red dots, specimens of *Nemertopsis caete*, sp. nov.; green dots, specimens of *Nemertopsis pamelaroeae*, sp. nov. Inset at top left shows the contribution of each PC and the numbers on top of each bar show the percentage contribution of each PC. An asterisk indicates less than 1% contribution.

DNA extraction, PCR amplification, and sequencing

Genomic DNA was extracted from 62 specimens using the Doyle and Doyle (1987)

protocol. Twenty-nine of the 62 individuals had the partial regions of *COI* and *16S* rRNA

amplified using the primer pairs LCO1490/HCO2198 (Folmer *et al*. 1994) and ar-L/br-H

(Palumbi *et al*. 1991), respectively. Polymerase chain reactions (PCR) were carried out using Multiplex PCR Plus Kit (Qiagen) or Time Hifi DNA Polymerase (ECRA biotec), in the following conditions: initial denaturation at 94°C for 1 min; 35 cycles of denaturation at 94°C for 30 s; annealing at 45–47°C (*COI*) or 44–50°C (*16S* rRNA) for 1 min; extension at 72°C for 1 min; and final extension at 68°C for 3 min. PCR products were purified either with illustra ExoProStar (GE Healthcare Life Sciences) or PEG8000 (protocol modified from http://labs.icb.ufmg.br/lbem/protocolos/peg.html). Purified products were sequenced at Myleus Facility (Belo Horizonte, MG-Brazil). Resulting sequences were trimmed to remove primer regions and low quality ends, and complementary strands proofread against each other using GeneStudioPro (GeneStudio, Inc.) and checked for contaminants using the BLASTN tool.

Mitochondrial data

Alignment was performed by the online version of MAFFT (ver. 7, https://mafft.cbrc.jp/alignment/server/; Katoh *et al*. 2019), using the Auto strategy and remaining default parameters. Sequences of *COI* and *16S* rRNA were first analysed separately using Automatic Barcoding Gap Discovery (ABGD) online software (ver. 1, https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html; Puillandre *et al*. 2012) with default values for all parameters, and the Jukes–Cantor (JC69) distance method. ABGD uses an algorithm that calculates the divergence between sequences and automatically infers the barcoding gap between groups of sequences. The sequences were also concatenated and this dataset was used as input for phylogenetic inferences in RAxML ver. 8.2.12 (Stamatakis 2014), as available in CIPRES (Miller *et al*. 2010), under the GTRGAMMA model with 1000 bootstraps. The GTRGAMMA is the most parameter-rich model, and phylogenetic inferences using this model lead to the same results as models selected by other criteria (Abadi *et al*. 2019). Sequences from *Nemertopsis tetraclitophila* Gibson, 1990 available in

GenBank (accessions KF572482 and NC024823) were used as root for this analysis. The resulting tree was visualised in Figtree (ver. 1.4.3, http://tree.bio.ed.ac.uk/software/figtree/; Rambaut 2014) and used as input for species delimitation analyses in the online version of the Bayesian implementation of the Poisson Tree Processes (bPTP) (Zhang *et al*. 2013), with default parameters. PTP is a tree-based method that uses the number of substitution events as given by branch lengths of an input phylogram to infer intra- and interspecific relationships between sequences. The bPTP version adds Bayesian support values to the delimited species in the input tree (Zhang *et al*. 2013). Concatenated sequences were also used to construct haplotype networks in PopArt (ver. 1.7, http://popart.otago.ac.nz/index.shtml; Leigh and Bryant 2015). Diversity indices were calculated in DnaSP (ver. 6, http://www.ub.edu/dnasp/index_v5.html; Rozas *et al*. 2017).

In order to better understand the relationship of our *Nemertopsis* species to others in the genus, we also combined all *Nemertopsis* sequences of *COI* and *16S* rRNA available in GenBank with our sequences to perform a phylogenetic inference. We also added *COI* and *16S* rRNA sequences of two *N. bivittata* from Buenos Aires, Argentina, one provided by Eduardo Zattara (GenBank accessions MW074335/MW074169) and another previously available in GenBank (Accession MK047679/MK067305), and of one specimen from the north-west coast of Spain (Atlantic Cantabrian Sea) provided by Nuria Anadón (GenBank accessions MW065563/MW063537). The phylogenetic trees were inferred for each gene separately in RAxML (ver. 8.2.12, https://cme.h-its.org/exelixis/web/software/raxml/) under the GTRGAMMA model with 1000 bootstraps. The resulting trees were visualised in Figtree.

Library preparation, sequencing and SNPs prospection

Approximately 100 ng of the extracted genomic DNA per sample of 53 out of 62 specimens were sent to EcoMol Consultoria e Projetos Ltda (Brazil), where genotyping-by-sequencing (GBS) libraries were prepared according to the protocol of Elshire *et al*. (2011) and modification of Nunes *et al*. (2017). Briefly, DNA was digested with *Pst*I (De Donato *et al*. 2013), adapters were ligated, and fragments amplified. After amplification, fragments were quantified using the KAPA library quantification kit (Roche Sequencing Solutions) and fragment size checked using Bioanalyzer (Agilent Technologies). Finally, the libraries were sequenced in 100 bp single-end fragments using Illumina Hiseq 2500 platform, at the Centro de Genômica Funcional ESALQ-USP.

Genomic data

Sequenced fragments were first filtered using Seqyclean (ver. 1.10.09, https://github.com/ibest/seqyclean; Zhbannikov *et al*. 2017), removing sequences with average phred quality score ≤20, contaminants and adapters (based on the UniVec database; https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/). After filtering, SNP calling was performed using the *de novo* algorithm in Ipyrad (ver. 0.7.30, https://ipyrad.readthedocs.io/en/latest/; Eaton and Overcast 2020). Adaptors were trimmed, and bases with phred score below 33 were converted to N. Reads with more than five Ns or shorter than 35 bp were discarded. The minimum read depth was set to six for calling consensus sequences within samples and the maximum depth was set to 100 000. The clustering threshold was set to 90%. All loci sharing heterozygotic sites across more than eight individuals were filtered out to avoid paralogs, and the maximum number of SNPs per locus was set to 30. A locus had to be present in at least 20% of samples to be kept in the final dataset. To avoid linkage disequilibrium only one random SNP was called per site. All

other parameters were left at default values. To better visualise how the SNPs were distributed between the species, the final dataset was converted to an occupancy matrix using R software (ver. 3.5, R Core Team). This matrix was visualised in the Matrix Condenser software (ver. 1, https://bmedeiros.shinyapps.io/matrix_condenser/; de Medeiros and Farrell 2018) with the option divergent to better visualise the SNP distribution between groups. This tool plots a graph with the samples as rows and SNPs as columns, using the option divergent, generating a PCA to sort loci and samples, maximising the differences.

The SNP matrix was also used as input for phylogenetic analysis using IQ-TREE (ver. 1.5.6, http://www.iqtree.org/; Nguyen *et al*. 2015), under the GTR+ASC model, with 1000 bootstrap replicates. This model was chosen due to the lack of constant sites in the SNP data. The ascertainment bias correction prevents overestimating branch lengths (Lewis 2001). The consensus tree was visualised using FigTree.

Taxonomic results

Class **HOPLONEMERTEA** Hubrecht, 1879

Order **MONOSTILIFERA** Brinkmann, 1917

Suborder **EUMONOSTILIFERA** Chernyshev, 2003

Family **EMPLECTONEMATIDAE** Bürger, 1904

Genus *Nemertopsis* Bürger, 1895

Nemertopsis berthalutzae, sp. nov.
(Fig. 3)

Material examined

Thirteen adults, including holotype, from Balneário dos Trabalhadores beach. Three adults from Sabiaguaba beach. Two immature adults from Itaoca beach. Three immature adults from Mambucaba beach. Three adults from Prainha beach. One adult from Barra do Corumbê. Two adults from Pontal da Cruz pier. Two adults from Igararecê marina. Nineteen adults from Araçá beach. One adult from Ribeirão da Ilha beach. *COI* sequences from 12 adults (GenBank accession numbers in Table S1) as well as 2903 SNPs from 19 specimens. Twenty individuals remain as ethanol-preserved specimens or extracted DNA at Laboratório de Diversidade Genômica, IB-USP.

Type material

Type material deposited at the Museum of Zoology of the University of São Paulo (MZUSP) as ethanol preserved specimens, head previously fixed in formaldehyde, (MZUSP30 (holotype), MZUSP31 and MZUSP32 (two paratypes, directly fixed in 99% ethanol)). The partial COI sequence designed as topogenetype is deposited at GenBank accession MT433976.

Description

External appearance. Body length 8–30 mm, width 0.5–1 mm. Relatively translucent whitish body wall, yellowish internally, general appearance sometimes grey to brown in posterior region (Fig. 3*A*). Head round, wider than body, without internal white cap. One transverse cephalic furrow visible in ventral region, in front of cerebral ganglia. Small ventral rhynchostomopore, only visible in living animals. Four dark red simple precerebral eyes of similar size, arranged approximately as corners of a square. Two solid longitudinal lines along dorsal surface, dark brown, more reddish near posterior of body; extend beyond anterior eyes, meet at body posterior in most specimens. (Fig. 3*A*, *B*); lines more reddish

posteriorly, meet at posterior tip (Fig. 3*A*). Lines can have variable width sometimes merging in posterior region (Fig. 3*A*). Dark dots randomly distributed along entire body between parallel lines. Posterior tip of body tapered without any cirri.

Internal features.   Translucent cerebral ganglia. Small cerebral organs. Short rhynchocoel reaching to around first third of the body. Transparent anterior proboscis. Posterior proboscis opaque. Diameter of posterior third of proboscis less than 50% of body diameter. Central stylet straight (17–31 µm), smooth, conical shape. Basis mostly oval, sometimes cylindrical (25–55 µm), mostly without waist, but sometimes waist present. Basis round or sometimes flat posteriorly (Fig. 3*D*–*F*). Basis length/width ratio 2.9–4.9. Basis/stylet length ratio 1.2–2. Two stylet pouches, lateral to central stylet, 3–4 accessory stylets each (Fig. 3*D*, *E*). Lateral intestinal diverticula beginning at end of rhynchocoel, extending to posterior tip of body. Separate sexes. Transient serial white gonads arranged between intestinal diverticula in both sexes. Mature females have ~10 eggs per ovary (Fig. 5*C*).

Distribution

The species has been collected in the north-east, south-east and south coasts of Brazil in the states Ceará, Espiríto Santo, Rio de Janeiro, São Paulo, and Santa Catarina. The species was also collected by Dr. Eduardo Zattara in Mar Chiquita, Buenos Aires, Argentina in 2013. The type locality is Balneário dos Trabalhadores in São Paulo state.
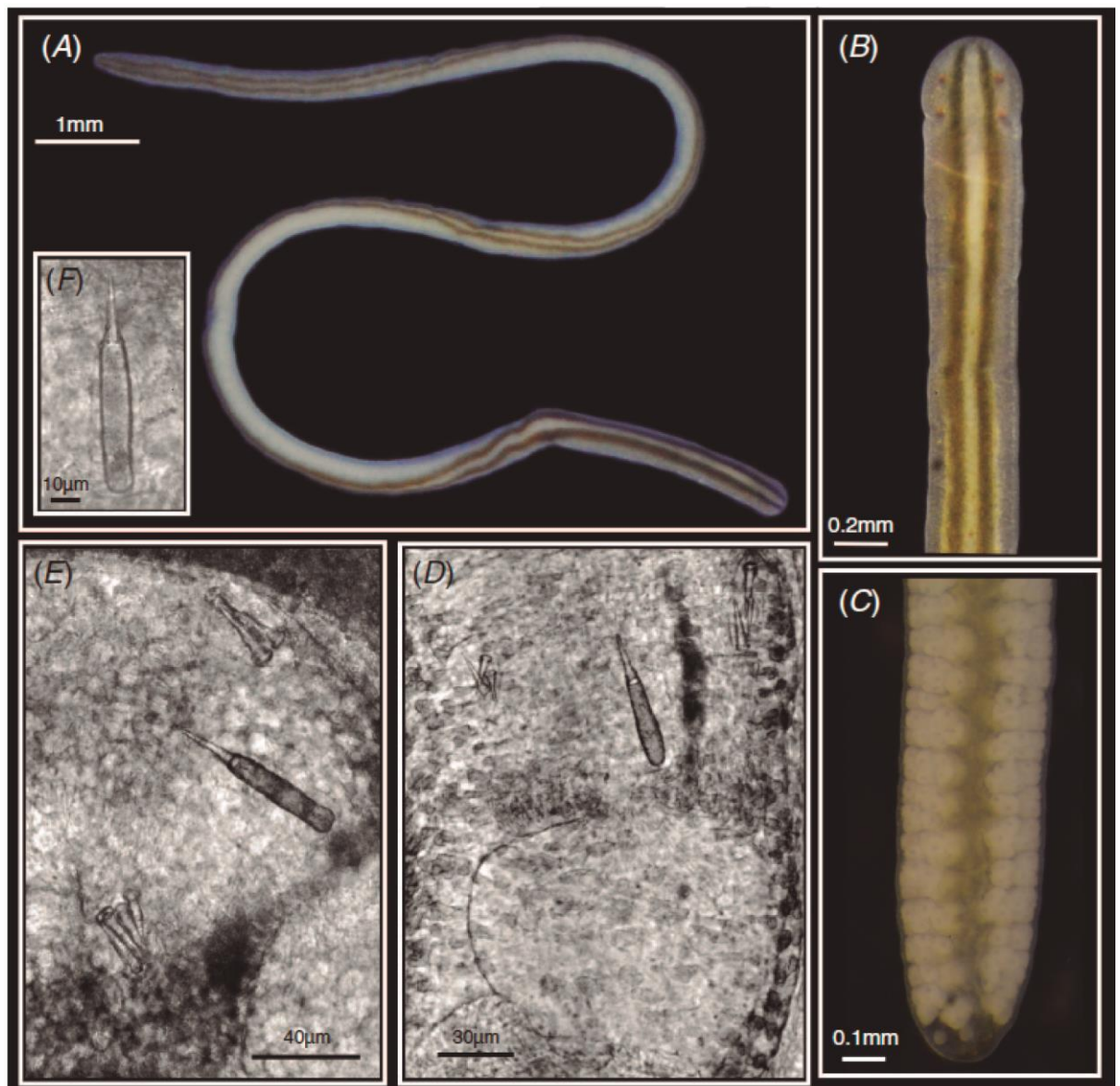
Figure 3. Photomicrographs of *Nemertopsis berthalutzae*, sp. nov. (*A*) Live specimen in dorsal view. (*B*) Detail of head. (*C*) Detail of posterior end of a female, showing gonads in development. (*D*) Detail of central stylet and stylet pouches showing a basis in conical shape without waist and round end. (*E*) Detail of central stylet and stylet pouches showing a basis in cylindrical shape with waist and flat end. (*F*) Detail of central stylet showing a basis in cylindrical shape without waist and flat end.

Remarks

So far, this is the only species of the three species here described known from the southern region of South America. This species is most similar to *Nemertopsis bivittata* (=*peronea*) from the Mediterranean type locality in having a pair of dorsal stripes reaching the anterior cephalic margin and not touching each other anteriorly (Bürger 1895; Corrêa 1955). It differs

in that Bürger (1895) states that the Mediterranean form has a stylet longer than the basis, though that statement is not supported by his figure (Bürger 1895: pl. 8, fig. 9), which shows them as approximately equal. The *N. bivittata COI* sequence from the Atlantic Cantabrian Sea has a pairwise distance ranging from 0.044 to 0.11 to *N. berthalutzae* sequences (Table 2). *Nemertopsis berthalutzae* can be differentiated from other species of the complex in Brazil by the lack of a transversal line in the head and longitudinal lines reaching the anterior tip (Table 3). The species also differ in *COI* sequences (pairwise distance ranges from 0.15 to 0.18 in *N. caete* and from 0.14 to 0.17 in *N. pamelaroeae*: Table 2). Differences between *N. berthalutzae* and other *Nemertopsis* species with dark dorsal lines are summarised in Table 3. *Nemertopsis berthalutzae* lives in oyster beds of *Crassostrea brasiliana* (Lamarck, 1819) and *Crassostrea mangle* Amaral & Simone, 2014, and in barnacle beds of *Tetraclita* Schumacher, 1817, *Balanus* Costa, 1778 and *Amphibalanus* Pitombo, 2004, on natural and artificial substrata, even with high human impact. At least in the south-east coast of Brazil, this species reaches its reproductive peak in late January, but no reproductive specimens were collected on the north-east coast. When exposed to 7.5% $MgCl_2$, specimens tend to produce excess mucus, sometimes contracting by coiling into a spiral.

Etymology

The specific name is a noun in the genitive case, after the Brazilian zoologist and feminist activist Bertha Lutz. She was an activist for women's right to vote, to work outside the home, and for earnings equity.

Nemertopsis caete, sp. nov.

(Fig. 4)

Material examined

Twenty-three adults, including holotype and paratypes, and 15 newly settled juveniles all collected in oyster beds at Praia do Prado. One adult from Sabiaguaba beach. *COI* and *16S* rRNA sequences from five adults (GenBank accession numbers in Table S1) as well as 2903 SNPs from 17 specimens were analysed. Twenty-one individuals remain as ethanol-preserved specimens or extracted DNA at Laboratório de Diversidade Genômica, IB-USP.

Type material

Type material deposited at the Museum of Zoology of the University of São Paulo (MZUSP) as ethanol-preserved specimens: MZUSP24 (holotype), MZUSP25 and MZUSP26 (paratypes). The partial COI and 16S rRNA sequences designated as topogenetypes are recorded in GenBank accession MT433994 and MT434998 respectively.

Description

External appearance. Body length 3–46 mm, width 0.5–0.8 mm. Relatively translucent whitish body wall, yellowish internally, paler in posterior region (Fig. 4*A*). Head rounded, wider than body, with indistinct internal white cap. One transverse cephalic furrow visible in ventral region, in front of cerebral ganglia. Small ventral rhynchostomopore, visible only in living animals. Four dark, simple precerebral eyes similar in size, arranged approximately as a square; visible from ventral side, especially in compressed specimens. Two longitudinal lines on dorsal surface, dark brown to more reddish near posterior; meet at body posterior, extend to anterior eyes with complete or medially broken transverse line passing over each eye (Fig. 4*B*, *C*). Both parallel and transverse lines can be paler in younger animals, and almost absent in the posterior region of newly settled juveniles. Dark dots randomly distributed along entire body between parallel lines, observed only in larger specimens. Posterior tip of body tapered, without cirri.

Internal features. Based on 12 specimens compressed under coverslip. Translucent cerebral ganglia. Small cerebral organs. Short rhynchocoel, reaching to around first third of the body. Transparent anterior proboscis. Posterior proboscis opaque. Diameter of posterior third of proboscis less than 50% of body diameter. Central stylet straight (35–27 µm), smooth and conical shape. Basis cylindrical (45–46 µm), no waist, posterior flat (Fig. 4D). Basis length/width ratio 3.6–5. Basis/stylet length ratio 1.3–1.7. Two stylet pouches, lateral to central stylet, 3–4 accessory stylets each (Fig. 4D). Lateral intestinal diverticula beginning at posterior of rhynchocoel, present to posterior tip of body. Separate sexes. Transient serial white gonads arranged between intestinal diverticula in both sexes.

Distribution

So far known only from two localities, both in north-eastern Brazil: the type locality, Prado beach, Maceió (Alagoas, Brazil), and Sabiaguaba beach, Fortaleza (Ceará, Brazil).

Remarks

This species can be differentiated from other species of the *Nemertopsis bivittata* complex in Brazil by a transverse bar (complete or medially broken) that always extends across the anterior ends of the two longitudinal lines to reach the first pair of eyes. They also can be differentiated by *COI* and *16S* rRNA sequences (pairwise distance for *COI* ranges from 0.15 to 0.18 for *N. berthalutzae* and from 0.005 to 0.14 for *N. pamelaroeae*; pairwise distance for *16S* rRNA ranges from 0.008 to 0.12 for *N. pamelaroeae*: Table 2). Differences between *Nemertopsis caete* and other *Nemertopsis* species are summarised in Table 3. The species lives in oyster beds of *Crassostrea mangle* and *Crassostrea brasiliana* and among barnacles of *Tetraclita* sp. and *Amphibalanus* sp., in both natural and artificial environments. Reproduction seems to occur through summer (December–March), since many newly settled individuals were collected in late January, as well as some mature adults. When exposed to

7.5% MgCl$_2$, specimens tend to produce excess mucus and contract without coiling into a spiral.

Etymology

The specific name is a tribute to the indigenous tribe Caeté, which used to live around the region where most specimens were collected, until ~1600 CE, when their tribe effectively became extinct at the hands of the Portuguese colonisers.
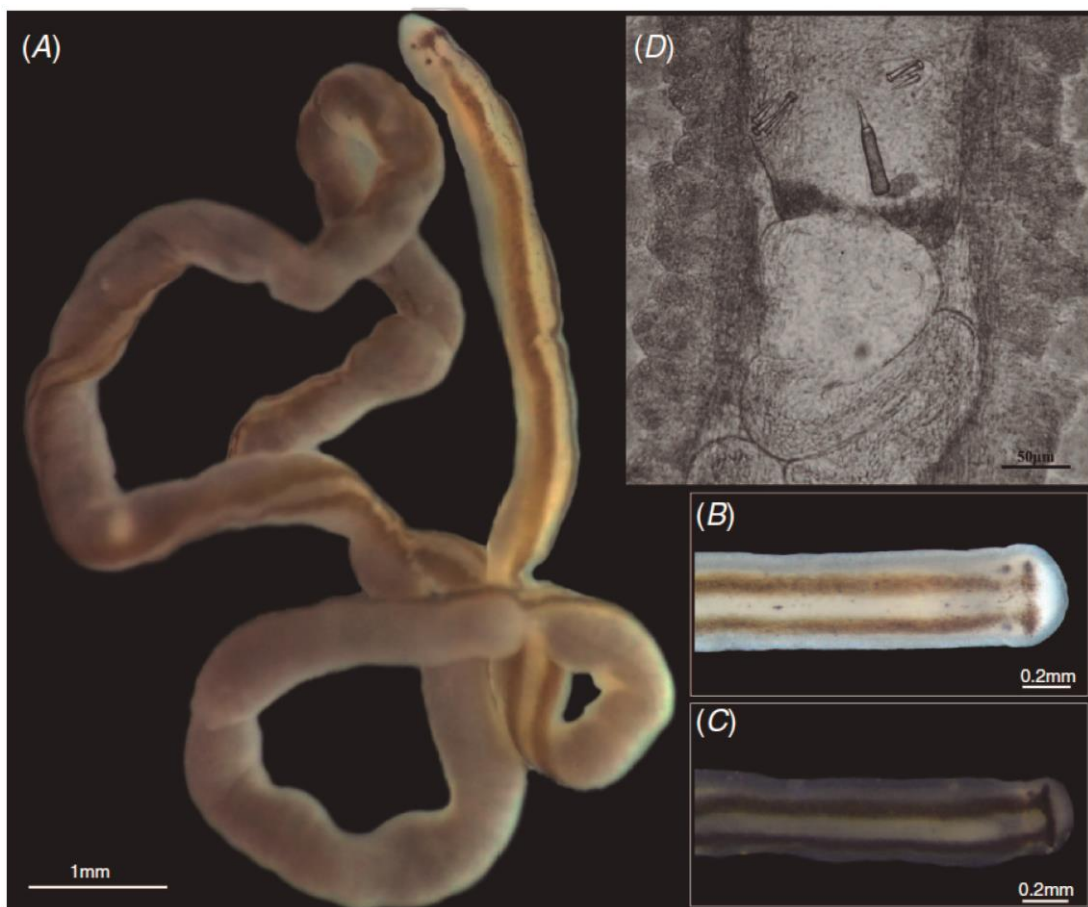
Figure 4. Photomicrographs of *Nemertopsis caete*, sp. nov. (*A*) Live specimen in latero/dorsal view. (*B*) Detail of head of a specimen with dashed transverse line. (*C*) Detail of head of a specimen with solid transverse line. (*D*) Detail of central stylet and stylet pouches seen under light microscopy.

Nemertopsis pamelaroeae, sp. nov.

(Fig. 5)

Material examined

Thirty-nine adults, including holotype and paratypes, from Jabaquara Beach. Two adults from Pedra Rachada Beach. Two adults from Itaipu Beach. Four adults from Igararecê marina. *COI* and *16S* rRNA sequences from six adults (GenBank accession numbers in Table S1) as well as 2903 SNPs from 16 specimens. Twenty individuals remain as ethanol-preserved specimens or extracted DNA at Laboratório de Diversidade Genômica, IB-USP.

Type material

Type material deposited at the Museum of Zoology of the University of São Paulo (MZUSP) as ethanol-preserved specimens, head previously fixed in formaldehyde: MZUSP27 (holotype), MZUSP28 and MZUSP29 (two paratypes directly fixed in 99% ethanol). The partial COI and 16S rRNA sequences designated as topogenetypes are recorded in GenBank, accessions MT433997 and MT435005 respectively.

Description

External appearance. Body length 9–51 mm, width 0.5–1 mm. Relatively translucent whitish body wall, yellowish internally (Fig. 5*A*). Head round, wider than body, with distinct internal white cap. One transverse cephalic furrow visible in ventral region, anterior to cerebral ganglia. Small ventral rhynchostomopore, visible only in living animals. Four dark red, simple precerebral eyes of similar size, arranged approximately as a square. Two solid longitudinal lines along dorsal surface, dark brown, more reddish near posterior of body; extend to anterior eyes, meet at body posterior. Solid transverse bar at anterior pair of eyes unites longitudinal lines but does not extend laterally beyond them (Fig. 5*A*, *B*). Dark dots randomly distributed along entire body between parallel lines. Posterior tip of body tapered, without cirri.

Internal features. Based on 39 specimens compressed under coverslip. Translucent cerebral ganglia. Small cerebral organs. Short rhynchocoel reaching to around first third of the body. Transparent anterior proboscis. Posterior proboscis opaque. Diameter of posterior third of proboscis less than 50% of body diameter. Central stylet straight (17–40 μm), smooth and conical shape. Basis cylindrical (26–70 μm), usually with waist, but sometimes waist not present. Basis posterior flat or sometimes concave (Fig. 5*C–E*). Basis length/width ratio 3.1–6.2. Basis/stylet length ratio 1.5–2.2. Two stylet pouches, lateral to central stylet, 3–6 accessory stylets each (Fig. 5*C*, *D*). Lateral intestinal diverticula beginning at posterior end of rhynchocoel, extending to posterior end of body. Separate sexes. Transient serial white gonads arranged between intestinal diverticula in both sexes. Mature females have ~10 eggs per ovary.

Distribution

The species was collected on the north-east and south-east coasts of Brazil in the states of Ceará, Espiríto Santo, Rio de Janeiro and São Paulo. One specimen of *N. bivittata* from Fort Pierce (Florida), collected by Dr. Eduardo Zattara also seems to belong to this species. Type locality is Praia do Jabaquara in Rio de Janeiro state.

Remarks

This species can be differentiated from other species of the *Nemertopsis bivittata* complex in Brazil by the transverse bar that never extends laterally beyond the longitudinal lines to the anterior eyes. They also can be differentiated by *COI* and *16S* rRNA sequences (pairwise distance for *COI* ranges from 0.14 to 0.17 for *N. berthalutzae* and from 0.005 to 0.14 for *N. caete*; pairwise distance for *16S* rRNA ranges from 0.008 to 0.12 for *N. caete*: Table 2). Morphological differences between *Nemertopsis pamelaroeae* and other *Nemertopsis* species with longitudinal dorsal lines are summarised in Table 3. *Nemertopsis pamelaroeae* lives in

oyster beds of *Crassostrea brasiliana* and *Crassostrea mangle* and in barnacle beds of *Tetraclita* sp., *Balanus* sp. and *Amphibalanus* sp., on natural and artificial substrata, even with high human impact. On the south-east coast of Brazil, this species reaches its reproductive peak in late January, but no reproductive specimens were collected on the north-east coast. Sequences from *Tetraclita* sp. and *Amphibalanus* sp. were retrieved from some specimens, which indicates that this species feeds upon different species of barnacles. When exposed to 7.5% $MgCl_2$, specimens tend to produce excess mucus and contract into a spiral.

Etymology

The specific name is a noun in genitive case, after the great nemertean specialist Dr Pamela Roe, who passed in 2018. She made important contributions to the knowledge of nemertean biology and dedicated great effort to science education of disadvantaged undergraduate and K-12 students.
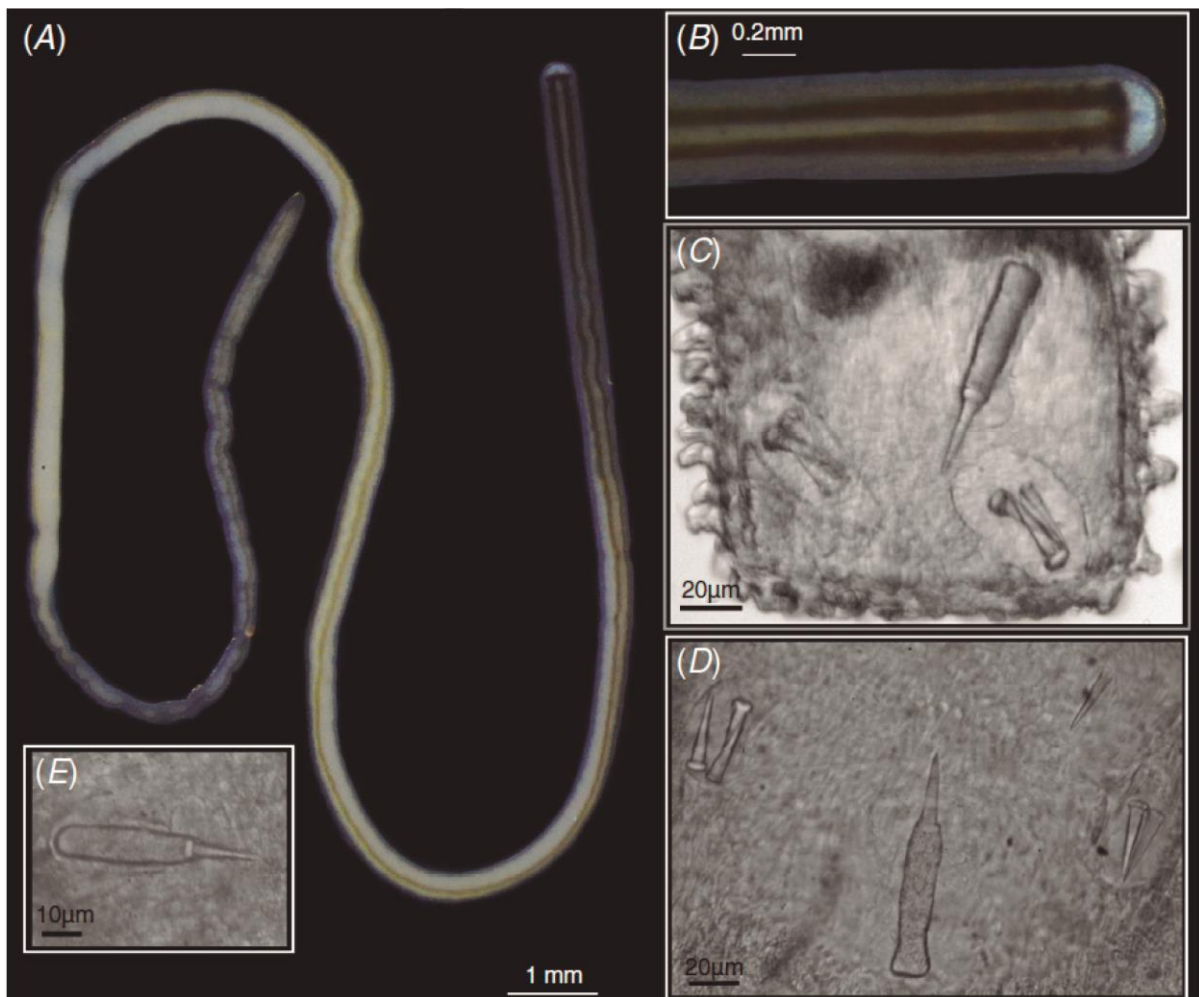
Figure 5. Photomicrographs of *Nemertopsis pamelaroeae*, sp. nov. (*A*) Live specimen in dorsal view. (*B*) Detail of head. (*C*) Detail of central stylet and stylet pouches showing a basis in conical shape without waist and flat end. (*D*) Detail of central stylet and stylet pouches showing a basis in cylindrical shape with waist and flat end. (*E*) Detail of central stylet showing a basis in cylindrical shape without waist and round end.

Genetic Principal component analysis

PCA of external features clusters together individuals of *N. caete* and *N. pamelaroeae*. All individuals of *N. berthalutzae* are found in a separate cluster (Fig. 2*B*), which is explained by the PC1 axis accounting for 65% of the variation. The first principal component is influenced by all characters, but mostly by the position of the end of the dorsal stripes in the head, and the distance between the stripes. The second is influenced similarly by all characters.

Table 2. Inter- and intraspecific uncorrected p-distances for *16S* rRNA and *COI* (bold text) gene fragments in *Nemertopsis* species from Brazil

	N. bivittata (Cantabrian Sea)	*N. berthalutzae*	*N. caete*	*N. pamelaroeae*
Nemertopsis bivittata (Cantabrian Sea)	0			
	0			
Nemertopsis berthalutzae	0.011–0.012	0		
	0.044–0.11	**0.02**		
Nemertopsis caete	0.12	0.13	0	
	0.14–0.15	**0.15–0.18**	**0.01**	
Nemertopsis pamelaroeae	0.12	0.12	0.008–0.12	0.03
	0.14–0.15	**0.14–0.17**	**0.005–0.14**	**0.1**

Molecular markers

GenBank accession numbers of *16S* rRNA and *COI* sequences are presented in Table S1. We obtained *16S* rRNA sequences for five *Nemertopsis caete* and six *Nemertopsis pamelaroeae* but none for *Nemertopsis berthalutzae*. These sequences are 475 bp long, with 57 variable sites, all being parsimony informative. We obtained *COI* sequences for five *N. caete*, two *N. pamelaroeae* and 17 *N. berthalutzae*. These sequences are 589 bp long, with 160 variable sites, of which 116 are parsimony informative and 44 are singletons.

Illumina sequencing resulted in 153 449 945 reads of 100 bp average length before filtering (Table S2) for all 53 specimens. The reads are deposited as the SRA Bioproject PRJNA629421. From those, 153 432 168 passed the quality filters for size and contamination; 2 011 579 were clustered per similarity; 535 216 of the clustered reads passed the depth filter, and 524 527 clustered reads passed the heterozygosity filter and were used in the SNP prospection step in Ipyrad (Table S2) with all individuals. During SNPs prospection, we used a threshold of 80% missing data per locus, due to the high divergence between species. Therefore, loci were retained only if present in at least 20% of samples, resulting in 2903 SNPs (Supplemental NEX file). From those, only 58 are shared between all three

species, 1127 are exclusive to *N. berthalutzae*, 1288 are exclusive to *N. caete* and 430 are exclusive to *N. pamelaroeae* (Fig. 6).

Mitochondrial diversity

Haplotype networks of all species are organised by federative states as a way to better represent the geographic patterns. The haplotype network of *N. berthalutzae* specimens shows higher haplotype and nucleotide diversity (Fig. 7*A*, Table 4); however, it does not indicate any geographic pattern. *Nemertopsis caete* shows low haplotype and nucleotide diversity (Fig. 7*B*, Table 4). The network of *N. pamelaroeae* also shows low haplotype and nucleotide diversity (Fig. 7*C*, Table 4), but one divergent specimen from Ceará (PNJ1) had several mutational steps from all the other specimens. This specimen is also in Clade A, along with specimens of *N. caete* on the phylogenetic trees (see below).

Species delimitation

Genomic and mitochondrial data separate the specimens into three main clades (Fig. 8, 9, S1). All specimens of *Nemertopsis caete* fall into a single clade in both analyses, separated as a long branch in the genomic analysis from the other two species (Fig. 8). The specimens of *Nemertopsis pamelaroeae* also form a clade in the genomic tree. However, in the mitochondrial tree one specimen from the north-east (PNJ1) clusters with specimens of *N. caete* that are geographically proximate. The specimens of *N. berthalutzae* form a clade in all phylogenetic hypotheses shown (Fig. 8, 9, S1). The three clades of both genomic and mitochondrial analyses correspond to the three clusters present in Matrix Condenser results (Fig. 6).
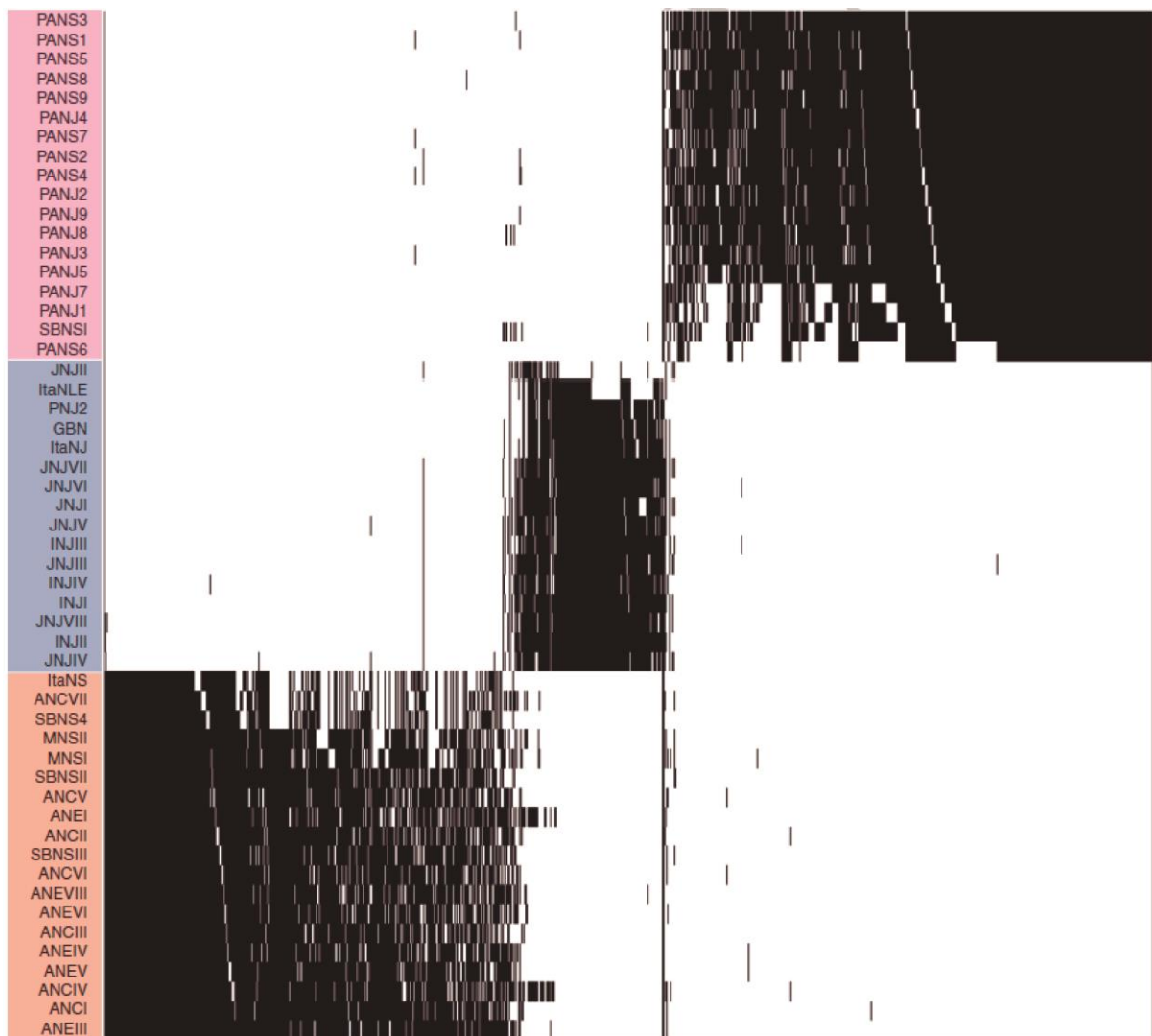
Figure 6. Matrix condenser results showing 2903 SNPs. Each column is a SNP, where the black bars represent a SNP present in the individual, and the blank bars indicate that the SNP is missing. Light orange shadow marks specimens of *Nemertopsis berthalutzae*, sp. nov.; pink shadow marks specimens of *Nemertopsis caete*, sp. nov.; blue shadow marks specimens of *Nemertopsis pamelaroeae*, sp. nov.
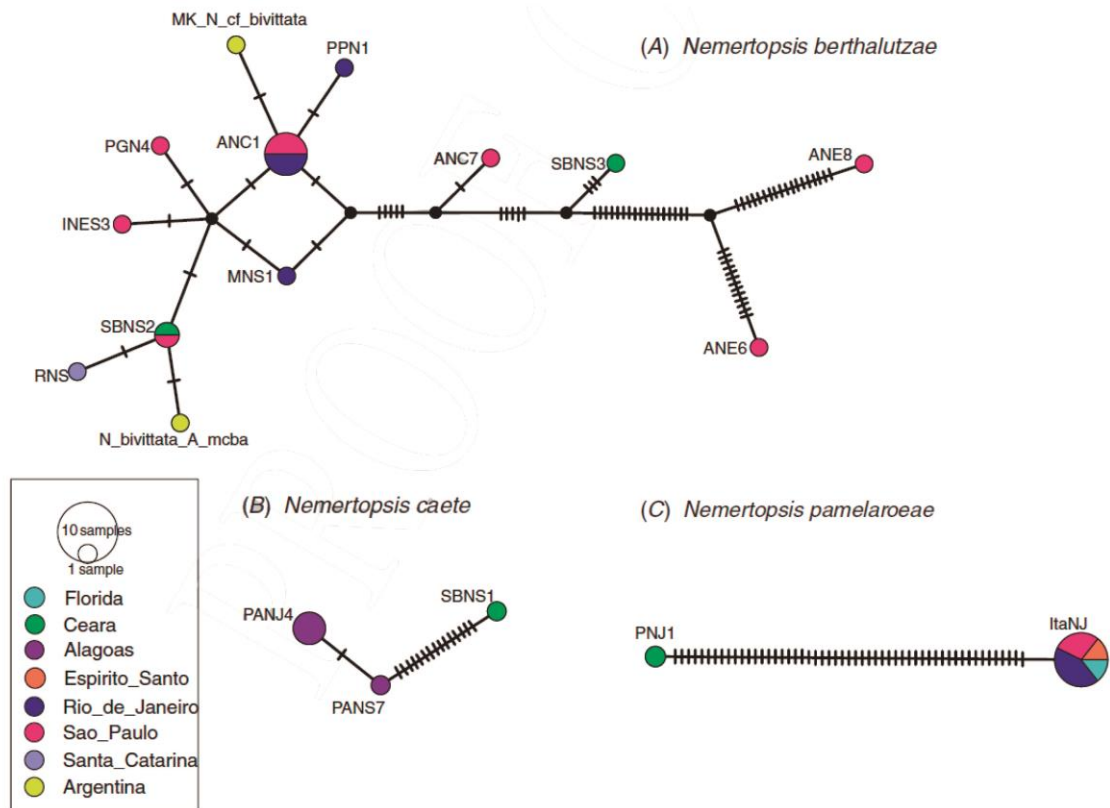
Figure 7. Haplotype network results using *COI* and *16S* rRNA sequences concatenated. (*A*) Haplotypes of *Nemertopsis berthalutzae*, sp. nov. (*B*) Haplotypes of *Nemertopsis caete*, sp. nov. (*C*) Haplotypes of *Nemertopsis pamelaroeae*, sp. nov. Colours of each federative state are according to Fig. 1.

Table 3. Morphological comparison between *Nemertopsis* species with dorsal lines

	N. berthalutzae sp. nov.	*N. bivittata*	*N. bullocki*	*N. caete* sp. nov.	*N. gracilis*	*N. mitellicola*	*N. pamelaroeae* sp. nov.	*N. quadripunctatus*	*N. tetraclitophila*
Size range	8–30 mm long, and 0.5–1.0 mm wide	8 mm long, and 0.5 mm wide	130–140 mm long, and 0.5–1.0 mm wide	3–46 mm long, and 0.5–0.8 mm wide	150 mm long, and less than 1 mm wide	Up to 150 mm long, and 0.4 mm wide	9–51 mm long, and 0.5–1.0 mm wide	1.5–80 mm long, and ~1 mm wide	4–35 mm long, and ~2 mm wide
Dorsal surface colour	Relatively translucent whitish body wall, yellowish internally, sometimes grey to brown in posterior region	White yellowish body	Pale yellow in the anterior region and blackish grey posteriorly	Relatively translucent whitish body wall, yellowish internally, paler in posterior region	Dull whitish with a tinge of brown, sometimes brownish. Much paler between the dorsal lines	Cream-white	Relatively translucent whitish body wall, yellowish internally	White or off-white, and very pale yellow between the dorsal lines	White or a very pale yellowish-white
Dorsal lines	Two longitudinal lines along dorsal surface, dark brown, more reddish near posterior of	Two dark longitudinal lines along dorsal surface, meet posteriorly; extend forward of anterior eyes; no	Two longitudinal, blackish brown lines, fade posteriorly to relatively darker background; meet posteriorly; connect	Two longitudinal lines on dorsal surface, dark brown to more reddish near posterior; meet at body	Two narrow longitudinal lines of deep brown meet posteriorly; extending through the body. On the head	Two dorsal dark-brown lines that fade posteriorly and do not meet; meet in a pigment patch on	Two longitudinal lines along dorsal surface, dark brown, more reddish near posterior of body; meet at body posteriorly;	Two longitudinal blackish-brown lines, but sometimes paler or not well demarcated in young specimens; usually meet at the tip of the	Two broad longitudinal dark-green lines along dorsal surface; meet posteriorly; meet anteriorly in green shield-shaped patch on head

	body; meet posteriorly in most specimens; extend forward of anterior eyes; no transversal line on the head	transversal line in the head	anteriorly by broad transverse band of the same colour anteriorly	posteriorly; extend to anterior eyes with complete or medially broken transverse line passing over each eye	they lie just internal to the eyes, do not extend to tip of head	the head	meet in solid transverse bar at anterior pair eyes that does not extend laterally beyond them	head without a transverse bar		
Number of eyes	Two pairs of eyes, arranged as corners of a square	Two pairs of eyes, arranged as corners of a square	Two pairs of eyes, arranged as corners of a square immediately in front of the red brain	Two pairs of eyes, arranged as corners of a square. First pair is covered by the transverse line	Two pairs of eyes, arranged as corners of a square	Two pairs of eyes, arranged as corners of a square	Two pairs of eyes, arranged as corners of a trapezium	Two pairs of eyes, arranged as corners of a square	Typically, two pairs of eyes, sometimes fragmented, appearing as up to four pair	No eyes evident in life
Basis/stylet length ratio	1.2–2.0	Bürger states stylet is longer than basis, but figure shows similar lengths	Approximately same length	1.3–1.7	Coe does not report lengths	1.75	1.5–2.2		3.5–2.7	4.0
Central	Central	Central	Central stylet	Central	Central	Central	Central	Central	Central stylet	Central stylet

stylet and stylet pouches	stylet straight (17–31 µm), smooth. Two stylet pouches, with 3–4 accessory stylets each	stylet straight. Two lateral stylet pouches.	straight, smooth, up to 59 µm. Two lateral stylet pouches.	stylet straight and smooth (35–27 µm). Two lateral stylet pouches with 3–4 accessory stylets each	stylet straight and smooth. Two lateral stylet pouches with 4–6 accessory stylets each	stylet spirally sculptured, up to 57 µm. Two lateral stylet pouches with 2–3 accessory stylets each	stylet straight and smooth (17–40 µm). Two lateral stylet pouches with 3–6 accessory stylets each	straight and smooth (17–22 µm). Two lateral stylet pouches with up to 5 accessory stylets each	straight and smooth (up to 30 µm). Four stylet pouches, with up to 6 accessory stylets each
Habitat	Free-living in oysters and barnacle beds	Free-living among *Mytilus* sp., barnacles and algae	Free-living, in rock crevices and in coarse sand	Free-living in oysters and barnacle beds	Free-living among mussels and other growths on rocks at low water	Lives in association with the gooseneck barnacle, *Capitulum mitella*	Free-living in oysters and barnacle beds	Lives in association with the gooseneck barnacle, *Capitulum mitella*, where it feeds upon the barnacle eggs	Lives in mantle cavity of the acorn barnacle *Tetraclita squamosa*

Figure 8. Tree resulting from maximum likelihood analysis of 2903 SNPs on IQ-TREE. Specimens of *Nemertopsis berthalutzae*, sp. nov. are shaded in light orange; specimens of *Nemertopsis caete*, sp. nov. are shaded in pink; and specimens of *Nemertopsis pamelaroeae*, sp. nov. are shaded in blue. Bootstrap support values ≥50 are shown on each branch.
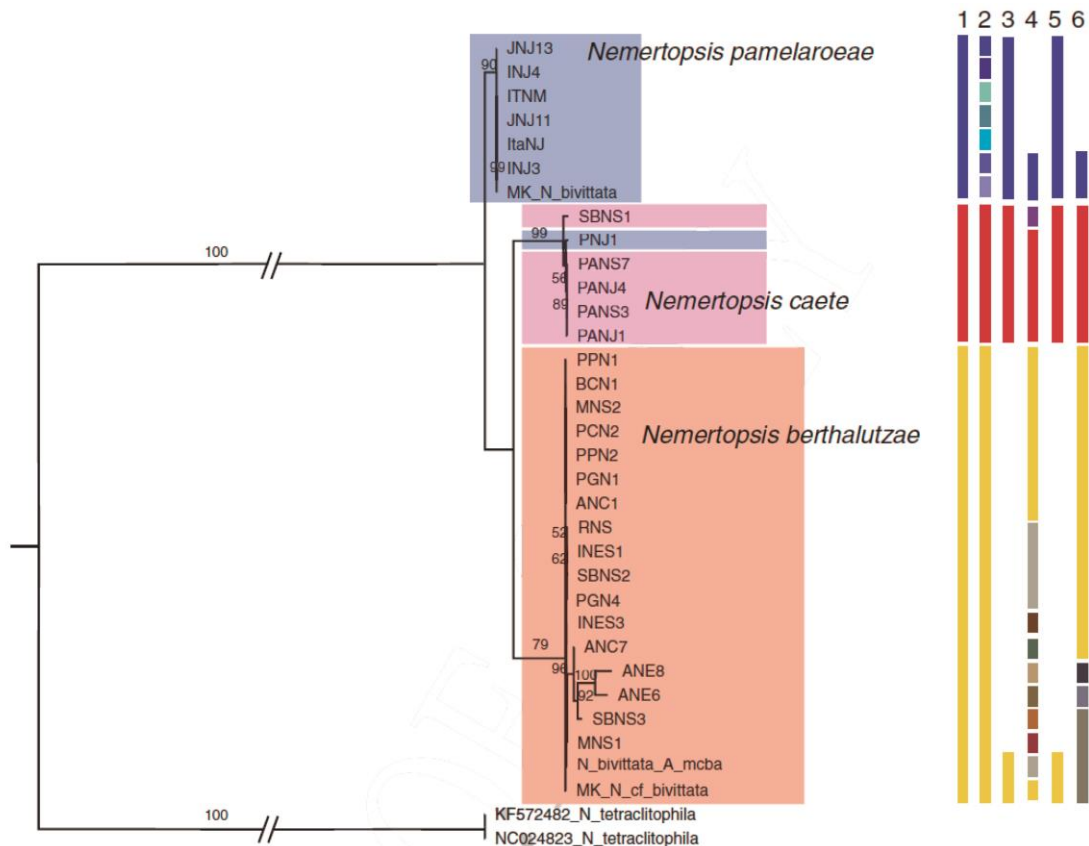
Figure 9. Tree resulting from maximum likelihood analysis of concatenated *COI* and *16S* rRNA fragments on RAxML (lnL = –3876.3318). *Nemertopsis berthalutzae*, sp. nov. is shaded in light orange; *Nemertopsis caete*, sp. nov. is shaded in pink; *Nemertopsis pamelaroeae*, sp. nov. is shaded in blue. Note: one specimen of *N. pamelaroeae* (PNJ1) is in the same clade as *N. caete* specimens. Tree is rooted by *Nemertopsis tetraclitophila* sequences. Bar 1 represents bPTP delimitation results of the maximum likelihood solution. Bar 2 represents bPTP delimitation results of the Bayesian solution. Bar 3 represents bPTP delimitation results of the maximum likelihood solution for *16S* rRNA. Bar 4 represents bPTP delimitation results of the maximum likelihood solution for *COI*. Bar 5 represents ABGD delimitation results using *16S* rRNA sequences. Bar 6 represents ABGD delimitation results using *COI* sequences. Bars with similar colours depict the same group. Blank spaces are a result of missing samples.

GenBank has previous *COI* and *16S* rRNA sequence data (GenBank accessions MK047680, KX377861, AJ436908, MK047679 for *COI* and AJ436798, MK067304, MK067305, JF277609 for *16S* rRNA) for four specimens identified as *Nemertopsis bivittata*. These sequences differ significantly from each other, as can be seen in phylogenetic analyses based on mitochondrial *16S* rRNA and *COI* data for our specimens and all *Nemertopsis* sequences previously available in GenBank (Fig. S1).

79

These analyses also do not show signs of any geographical pattern, with samples from nearby locations in different clades (Fig. S1). One sequence from Fort Pierce (Florida) clusters within the *N. pamelaroeae* clade, which expands this species' distribution to the North Atlantic Ocean. Sequences of two specimens from Argentina cluster with sequences of *N. berthalutzae*, which is the only one of the three species described here that was collected in South Brazil. However, three specimens identified as *Nemertopsis bivittata*, one from Colombia (GenBank accession KX377861) and two from the Atlantic coast of the USA (Florida and South Carolina; GenBank accessions AJ436908 and HQ848608), do not fall within the clades of the new species. The tree resulting from *16S* rRNA sequences shows *N. tetraclitophila* as sister to all other *Nemertopsis* species that belong to the nominal *Nemertopsis bivittata* complex (Fig. S1*A*). *COI* results place the clade formed by *N. berthalutzae* specimens as a sister to the clade comprising a specimen from north-west Atlantic Spain (GenBank MW065563) and another from South Carolina, USA (GenBank HQ848608). The two clades together form a clade closer to *Nemertopsis tetraclitophila* than to any of the three other *N. bivittata* specimens (Fig. S1*B*).

Table 4.   Total or average diversity indices of *Nemertopsis* species. H, haplotype diversity; S, number of parsimony informative sites; π, nucleotide diversity index

Species (fragment)	*H*	S	π
Nemertopsis caete (*16S* rRNA)	0	1	0
Nemertopsis caete (*COI*)	0.622	16	0.00978
Nemertopsis pamelaroeae (*16S* rRNA)	0.2857	0	0.03195
Nemertopsis berthalutzae (*COI*)	0.8592	33	0.03849

ABGD results from *COI* sequences show a barcoding gap at JC69 distance of 0.03–0.05, delimiting five groups. This barcoding gap clusters together most of the

specimens from *Nemertopsis berthalutzae*, including the two sequences from Argentina, but delimits two specimens from São Paulo (ANE6 and ANE8) as two different species. The two specimens of *Nemertopsis pamelaroeae* are also delimited as two different species: one (INJ3) as a species itself with the one sequence from Florida and the other clustered with the specimens of *Nemertopsis caete*, which all cluster as a single group (Fig. 9). The results from *16S* rRNA show a gap at JC69 distance of 0.015–0.12, separating three groups. The first comprises all specimens of *N. caete* and one of *N. pamelaroeae* from Paracuru (the same specimen with contradictory position in the phylogenies); the second comprises all other specimens of *N. pamelaroeae*; and the third has sequences of the two specimens from Argentina. Similar groups are delimitated in bPTP using the tree with the concatenated dataset. The maximum likelihood solution splits the tree into three major groups, matching the three species here described, except for the one contradictory specimen from the north-east (Fig. 9, S2). The Bayesian solution, however, splits the specimens of *N. pamelaroeae* as a different species each, and splits two specimens of *N. bertalutzae* (ANE6 and ANE8) as a different species each (Fig. 9, S2).

Discussion

The presence of cryptic species among different taxa is something already well known by many authors (Sundberg *et al*. 2009*a*; Adams *et al*. 2014; Hiebert and Maslakova 2015; Cornils *et al*. 2017; Fišer *et al*. 2018; Mendes *et al*. 2018; Quattrini *et al*. 2019). Discrete morphological differences may sometimes be present but, even when recognised, differences may be ambiguous; whether they represent infraspecific variation or species-level differences may not be objectively evident, even with phylogenetic analyses. Often, colour and colour pattern are the only external morphological variations among members of species groups of the phylum Nemertea

(Sundberg *et al*. 2009*b*, 2016*a*). A few nemertean specialists such as Bürger (1895) and Coe (1905) documented colour and pattern from living specimens in great detail and sometimes even used these as diagnostics, but most nemertean species have been described from preserved specimens that had lost most colour information by the time they were handled by a taxonomist. Hence, histological characterisation of internal anatomy was the main way to differentiate species and thereby led to the description of species that could not later be matched to specimens known from colour in life. Prior to the advent of molecular tools, colour and pattern variations, with no objective means to assess significance, were treated either as species or as infraspecific attributes, the latter sometimes used to characterise subspecies or varieties – e.g. *Nemertopsis gracilis bullocki* Coe, 1940. Molecular data provide independent, and perhaps more objective, means to delimit species, and Nemertea specialists increasingly have found colour pattern associated with genetic data useful to distinguish species (Sundberg *et al*. 2009*b*, 2016*a*; Hiebert and Maslakova 2015), though this is not always the case (Cherneva *et al*. 2019). Our molecular data corroborate morphological data and provide evidence for three evolutionary entities described here as three new species in the *Nemertopsis bivittata* complex that can be easily differentiated by colour pattern on the head (Fig. 3–5). Underestimating the significance of morphological variation can lead to a false inference that species have cosmopolitan distributions (Andrade *et al*. 2011; Mendes *et al*. 2018), which we suggest has been the case for *Nemertopsis bivittata sensu lato*, primarily because the cluster of basic external features that specimens of this group have in common – the pair of stripes, general colour, four eyes, body length and habitat – are relatively striking among nemerteans.

A basis/stylet length ratio of ~0.8–1.5 is common among monostiliferous nemerteans. Bürger (1895) describes the stylet of *Nemertopsis bivittata* (=*peronea*) from the type locality (Naples) as being longer than the basis but provides no measurements, whereas his fig. 8:9 shows the basis and stylet as approximately equal in length. Corrêa (1955) does not comment on basis/stylet ratio for specimens examined by her while in Naples. The range of 1.2–2.2 across the three new species here is small compared to ratios of 3 and 4 in *Nemertopsis quadripunctatus* and *Nemertopsis tetraclitophola* Gibson, 1990. The lack of differentiation between stylets (Fig. 3*D*, 4*C*, *D*, 5*E*, *F*) in the three Brazilian species may reflect ecological similarities, since the three species inhabit similar environments and feed upon similar food items (different species of acorn barnacles). Such similarity between cryptic species has been inferred for other invertebrates to be the result of either phylogenetic niche conservatism or morphological convergence (Fišer *et al*. 2018).

The histories of these cryptic species are as cryptic as the species themselves, as we know little about their biology, natural history or geographical distribution. However, the genomic data suggest that *Nemertopsis caete* has likely been isolated from the other two species for a long time, accumulating differences probably due to genetic drift (Som 2015). This isolation can be a result of various processes, such as speciation in a more distant location and recent introduction into the current collecting area, or speciation associated with novel niche exploitation leading to ecological isolation (Mayr 1947; Schluter 2001; Feder *et al*. 2005). Invasion by alien species is common in fouling communities, because such communities are easily transported on ship hulls or in ballast waters (Carlton 1985; Drake and Lodge 2004; Bailey 2015). Thus, wider sampling could reveal *N. caete* to have a distant geographical origin. The long branches in the genomic tree, however, can be a bias

caused by the large amount of missing data present in our dataset (Leaché *et al*. 2015). This missing data, in turn, most likely are a consequence of the divergence between species, as it is likely that the species do not share polymorphisms for the remaining SNPs.

The delimitation results show some discrepancies between mitochondrial and genomic data, with one individual of *Nemertopsis pamelaroeae* assigned to the *N. caete* clade in the mitochondrial results. This discrepancy could be either a sign of introgression between the species or incomplete lineage sorting in the mitochondrial tree (Rosenberg 2002; Sonsthagen *et al*. 2016; Bocek *et al*. 2019; Campbell *et al*. 2020). Introgression seems the most probable given that this mitochondrial haplotype is restricted (in this study) to geographically proximate regions (Ceará and Alagoas). That would indicate that reproductive isolation is not complete between these two species, which would be facilitated by both species having similar reproductive periods. Nonetheless, the species were never found at the same location. However, all three *Nemertopsis* species in Brazil have very patchy distributions; therefore, it is very possible that the location or locations where both species occur simply were not sampled.

The mitochondrial methods oversplit *N. pamelaroeae* and *N. berthalutzae*, with some individuals being depicted as different species. Such oversplitting has been reported previously for these tree-based species delimitation methods (e.g. Strong and Whelan 2019). It also is important to note that mitochondrial data in general, and especially the limited data available to us, have a lower power of resolution for phylogenies of closely related species in different taxa (Dupuis *et al*. 2012; Bray and Bocak 2016; Kobayashi and Sota 2019; Quattrini *et al*. 2019; Campbell *et al*. 2020). The mitochondrial markers can easily give mixed signals between the phylogenetic

and biogeographic history of a taxon, and confuse inferences, due to factors like adaptive introgression of mtDNA, demographic disparities and sex-biased asymmetries, or even hybrid zone movement as a result of human activities (Toews and Brelsford 2012, and references therein). The mitochondrial data, however, in conjunction with morphological traits can be used successfully to facilitate rapid species identification.

The use of genomic data to delimit and describe new species can improve our understanding of the natural history of a species, since these data can provide enough phylogenetic resolution to determine the relationship between closely related entities, even in recalcitrant taxa (Herrera and Shank 2016). Therefore, many studies have relied on this kind of data for species delimitation in taxa with histories of taxonomic confusion (e.g. Herrera and Shank 2016; Reyes-Velasco *et al*. 2018; Quattrini *et al*. 2019; Bocek *et al*. 2019; Stec *et al*. 2020).

The genus *Nemertopsis* contains 10 nominal species, with one species, *Nemertopsis gracilis*, having at least five variations identified (Norenburg *et al*. 2020), indicating the possibility of more cryptic species in the genus. To better understand their phylogenetic positions, the three new species of *Nemertopsis* were analysed in a broader context with other members of the genus for which *COI* or *16S* rRNA sequences were available in GenBank (Fig. S2). These trees suggest that other specimens identified as *Nemertopsis bivittata* at different locations may represent cryptic species as well, therefore emphasising the need for a broader revision of this species complex, including more sampling along the coasts of North America, Asia and especially the type locality of the species in the Mediterranean Sea.

In conclusion, delimiting cryptic species is more successful with a combination of mitochondrial, genomic and morphological data. However, mitochondrial results

need to be carefully evaluated, since they cannot fully resolve phylogeny of closely related species (Bray and Bocak 2016; Kobayashi and Sota 2019; Bocek *et al*. 2019; Campbell *et al*. 2020). Also, analyses with the use of just a few mitochondrial markers are more susceptible to error due to introgression or hybridisation between close species. This is why we believe that barcode regions should not be used as species delimitation tools, as Sundberg *et al*. (2016*b*) previously discussed, which we confirm here, for the first time, with genomic markers. We also highlight the importance of describing cryptic species as a way to improve integration between taxonomy and other biological sciences, as well as the conservation of such species.

Acknowledgements

References

Abadi, S., Azouri, D., Pupko, T., and Mayrose, I. (2019). Model selection may not be a mandatory step for phylogeny reconstruction. Nature communications 10(1), 1-11. doi 10.1038/s41467-019-08822-w

Adams, M., Raadik, T. A., Burridge, C. P., and Georges, A. (2014). Global biodiversity assessment and hyper-cryptic species complexes: more than one species of elephant in the room? *Systematic Biology* 63, 518–533. doi:10.1093/sysbio/syu017

Amaral, V. S., and Simone, L. R. L. (2014). Revision of genus *Crassostrea* (Bivalvia: Ostreidae) of Brazil. *Journal of the Marine Biological Association of the United Kingdom* 94(4), 811–836. doi:10.1017/S0025315414000058

Andrade, S. C. S., Norenburg, J. L., and Solferini, V. N. (2011). Worms without borders: genetic diversity patterns in four Brazilian *Ototyphlonemertes* species (Nemertea, Hoplonemertea). *Marine Biology* 158, 2109–2124. doi:10.1007/s00227-011-1718-3

Asch, R. G., Cheung, W. W. L., and Reygondeau, G. (2018). Future marine ecosystem drivers, biodiversity, and fisheries maximum catch potential in Pacific Island countries and territories under climate change. *Marine Policy* 88, 285–294. doi:10.1016/j.marpol.2017.08.015

Bailey, S. A. (2015). An overview of thirty years of research on ballast water as a vector for aquatic invasive species to freshwater and marine environments. *Aquatic Ecosystem Health & Management* 18, 261–268. doi:10.1080/14634988.2015.1027129

Bocek, M., Motyka, M., Kusy, D., and Bocak, L. (2019). Genomic and mitochondrial data identify different species boundaries in aposematically polymorphic *Eniclases* net-winged beetles (Coleoptera: Lycidae). *Insects* 10, 295. doi:10.3390/insects10090295

Bray, T. C., and Bocak, L. (2016). Slowly dispersing neotenic beetles can speciate on a penny coin and generate space-limited diversity in the tropical mountains. *Scientific Reports* 6(1), 33579. doi:10.1038/srep33579

Bürger, O. (1895). 'Die Nemertinen des Golfes von Neapel und der angrenzenden Meeres-Abschnitte.' (Verlag von R. Friedländer & Sohn.)

Campbell, E. O., Gage, E. V., Gage, R. V., and Sperling, F. A. H. (2020). Single nucleotide polymorphism-based species phylogeny of greater fritillary butterflies (Lepidoptera: Nymphalidae: Speyeria) demonstrates widespread mitonuclear discordance. *Systematic Entomology* 45, 269–280. doi:10.1111/syen.12393

Caplins, S., Norenburg, J. L., and Turbeville, J. M. (2012). Molecular and morphological variation in the barnacle predator *Nemertopsis bivitatta* [*sic*] (Nemertea, Hoplonemertea). *Integrative and Comparative Biology* 52, E24.

Carlton, J. T. (1985). Transoceanic and interoceanic dispersal of coastal marine organisms: the biology of ballast water. *Oceanography and Marine Biology* 23, 313–371.

Cherneva, I. A., Chernyshev, A. V., Ekimova, I. A., Polyakova, N. E., Schepetov, D. M., Turanov, S. V., Neretina, T. V., Chaban, E. M., and Malakhov, V. V. (2019). Species identity and genetic structure of nemerteans of the "Lineus ruber–viridis" complex (Muller, 1774) from Arctic waters. *Polar Biology* 42(3), 497–506. doi:10.1007/s00300-018-2438-7

Coe, W. R. (1905). Nemerteans of the west and northwest coasts of America. *Bulletin of the Museum of Comparative Zoology at Harvard College* 47, 1–318.

Coe, W. R. (1940). Revision of the nemertean fauna of the Pacific coasts of North, Central, and northern South America. *Allan Hancock Pacific Expeditions* 2, 247–322.

Cornils, A., Wend-Heckmann, B., and Held, C. (2017). Global phylogeography of *Oithona similis* sl (Crustacea, Copepoda, Oithonidae) – a cosmopolitan plankton species or a complex of cryptic lineages? *Molecular Phylogenetics and Evolution* 107, 473–485. doi:10.1016/j.ympev.2016.12.019

Corrêa, D. D. (1955). Os gêneros *Emplectonema* Stimpson e *Nemertopsis* Bürger (Hoplonemertini Monostilifera). *Boletim da Faculdade de filosofia ciências e letras. Universidade de São Paulo* 20, 67–86. doi:10.11606/issn.2526-3382.bffclzoologia.1955.120212

de Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., and Imumorin, I. G. (2013). Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLoS One* 8(5), e62137

de Lamarck, J. B. M. (1819). 'Histoire Naturelle des Animaux Sans Vertèbres. Tome Sixième, 1re Partie.' (Published by the Author: Paris, France.)

de Medeiros, B. A. S., and Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ* 6, e5089. doi:10.7717/peerj.5089

Dee, S. G., Torres, M. A., Martindale, R. C., Weiss, A., and DeLong, K. L. (2019). The future of reef ecosystems in the Gulf of Mexico: insights from coupled climate model simulations and ancient hot-house reefs. *Frontiers in Marine Science* 6, 691. doi:10.3389/fmars.2019.00691

Delle Chiaje, S. (1841). Descrizione e notomia degli animali invertebrati della Sicilia citeriore osservati vivi negli anni 1822–1830. *Napoli* 7, 86–173. doi:10.5962/bhl.title.10031

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19, 11–15.

Drake, J. M., and Lodge, D. M. (2004). Global hot spots of biological invasions: evaluating options for ballast-water management. *Proceedings of the Royal Society of London – B. Biological Sciences* 271, 575–580. doi:10.1098/rspb.2003.2629

Dryden, I. L., and Mardia, K. V. (1998). 'Statistical Shape Analysis.' (Wiley: Chichester, UK.)

Dupuis, J. R., Roe, A. D., and Sperling, F. A. (2012). Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology* 21(18), 4422–4436. doi:10.1111/j.1365-294X.2012.05642.x

Eaton, D. A. R., and Overcast, I. (2020). Ipyrad: interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36(8), 2592–2594. doi:10.1093/bioinformatics/btz966

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS)

approach for high diversity species. *PLoS One* 6, e19379. doi:10.1371/journal.pone.0019379

Feder, J. L., Xie, X., Rull, J., Velez, S., Forbes, A., Leung, B., Dambroski, H., Filchak, K. E., and Aluja, M. (2005). Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. *Proceedings of the National Academy of Sciences of the United States of America* 102, 6573–6580. doi:10.1073/pnas.0502099102

Fišer, C., Robinson, C. T., and Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* 27, 613–635. doi:10.1111/mec.14486

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3, 294–299.

Fowler, A. C. (2015). A simple thousand-year prognosis for oceanic and atmospheric carbon change. *Pure and Applied Geophysics* 172, 49–56. doi:10.1007/s00024-014-0892-x

Gibson, R. (1990). The macrobenthic nemertean fauna of Hong Kong. In 'Proceedings of the Second International Marine Biological Workshop: The Marine Flora and Fauna of Hong Kong and Southern China. The Marine Flora and Fauna of Hong Kong and Southern China II', 2-24 April 1986,Hong Kong. (Ed. B. Morton.) pp. 33–212. (Hong Kong University Press: Hong Kong.)

Gibson, R. (1995). Nemertean genera and species of the world: an annotated checklist of original names and description citations, synonyms, current taxonomic status, habitats and recorded zoogeographic distribution. *Journal of Natural History* 29, 271–561. doi:10.1080/00222939500770161

Herrera, S., and Shank, T. M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution* 100, 70–79. doi:10.1016/j.ympev.2016.03.010

Hiebert, T. C., and Maslakova, S. (2015). Integrative taxonomy of the *Micrura alaskensis* Coe, 1901 species complex (Nemertea: Heteronemertea), with

descriptions of a new genus *Maculaura* gen. nov. and four new species from the NE Pacific. *Zoological Science* 32, 615–637. doi:10.2108/zs150011

Junoy, J., Andrade, S. C. S., and Giribet, G. (2010). Phylogenetic placement of a new hoplonemertean species commensal on ascidians. *Invertebrate Systematics* 24, 616–629. doi:10.1071/IS10036

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics* 20, 1160–1166. doi:10.1093/bib/bbx108

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources* 11, 353–357. doi:10.1111/j.1755-0998.2010.02924.x

Kobayashi, T., and Sota, T. (2019). Divergent host use among cryptic species in the fungivorous ciid beetle *Octotemnus laminifrons* (Motschulsky, 1860), with descriptions of three new species from Japan. *Systematic Entomology* 44, 179–191. doi:10.1111/syen.12321

Leigh, J. W., and Bryant, D. (2015). Popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6, 1110–1116. doi:10.1111/2041-210X.12410

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., Durbin, R., Edwards, S. V., Forest, F., Gilbert, M. T. P., Goldstein, M. M., Grigoriev, I. V., Hackett, K. J., Haussler, D., Jarvis, E. D., Johnson, W. E., Patrinos, A., Richards, S., Castilla-Rubio, J. C., van Sluys, M. A., Soltis, P. S., Xu, X., Yang, H., and Zhang, G. (2018). Earth BioGenome Project: sequencing life for the future of life. *Proceedings of the National Academy of Sciences of the United States of America* 115, 4325–4333. doi:10.1073/pnas.1720115115

Lewis, P. O. (2001). A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic biology, 50(6), 913-925. doi: 10.1080/106351501753462876

Mayr, E. (1947). Ecological factors in speciation. *Evolution* 1, 263–288. doi:10.1111/j.1558-5646.1947.tb02723.x

Mendes, C. B., Norenburg, J. L., Solferini, V. N., and Andrade, S. C. S. (2018). Hidden diversity: phylogeography of genus *Ototyphlonemertes* Diesing, 1863 (Ototyphlonemertidae: Hoplonemertea) reveals cryptic species and high diversity in Chilean populations. *PLoS One* 13, e0195833. doi:10.1371/journal.pone.0195833

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 'Proceedings of the Gateway Computing Environments Workshop (GCE)', 14 November 2010, New Orleans, LA, USA'. INSPEC Accession Number 11705685, pp. 1–8. (IEEE.) doi:10.1109/GCE.2010.5676129

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32, 268–274. doi:10.1093/molbev/msu300

Norenburg, J. L., Gibson, R., Herrera-Bachiller, A., and Strand, M. (2020). World Nemertea Database. *Nemertopsis bivittata* (Delle Chiaje, 1841). In 'World Register of Marine Species'. Available at http://www.marinespecies.org/aphia.php?p=taxdetails&id=122715 [Verified 27 May 2020].

Nunes, J. R. S., Liu, S., Pértille, F., Perazza, C. A., Villela, P. M. S., de Almeida-Val, V. M. F., Hilsdorf, A. W. S., Liu, Z., and Coutinho, L. L. (2017). Large-scale SNP discovery and construction of a high-density genetic map of *Colossoma macropomum* through genotyping-by-sequencing. *Scientific Reports* 7, 46112. doi:10.1038/srep46112

Palumbi, S. R., Martin, A. P., Romano, S., Mcmillan, W., Stice, L., and Grabowski, G. (1991). The simple fools guide to PCR. Special publication, Department of Zoology, University of Hawaii, Honolulu, HI, USA.

Pitombo, F. B. (2004). Phylogenetic analysis of the Balanidae (Cirripedia, Balanomorpha). *Zoologica Scripta* 33(3), 261–276. doi:10.1111/j.0300-3256.2004.00145.x

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 21, 1864–1877. doi:10.1111/j.1365-294X.2011.05239.x

Quattrini, A. M., Wu, T., Soong, K., Jeng, M. S., Benayahu, Y., and McFadden, C. S. (2019). A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evolutionary Biology* 19(1), 116. doi:10.1186/s12862-019-1427-y

Reyes-Velasco, J., Manthey, J. D., Bourgeois, Y., Freilich, X., and Boissinot, S. (2018). Revisiting the phylogeography, demography and taxonomy of the frog genus *Ptychadena* in the Ethiopian highlands with the use of genome-wide SNP data. *PLoS One* 13(2), e0190440. doi:10.1371/journal.pone.0190440

Rosenberg, N. A. (2002). The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology* 61(2), 225–247. doi:10.1006/tpbi.2001.1568

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., and Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution* 34, 3299–3302. doi:10.1093/molbev/msx248

Sánchez, M. (1973). Sobre 4 especies de nemertinos de Quintero (Chile). *Studies on Neotropical Fauna* 8, 195–214. doi:10.1080/01650527309360462

Schluter, D. (2001). Ecology and the origin of species. *Trends in Ecology & Evolution* 16, 372–380. doi:10.1016/S0169-5347(01)02198-X

Schumacher, C. F. (1817). 'Essai d'un Nouveau Système des Habitations des Vers Testacés.' (Schultz: Copenhagen, Denmark.)

Som, A. (2015). Causes, consequences and solutions of phylogenetic incongruence. *Briefings in Bioinformatics* 16, 536–548. doi:10.1093/bib/bbu015

Sonsthagen, S. A., Wilson, R. E., Chesser, R. T., Pons, J. M., Crochet, P. A., Driskell, A., and Dove, C. (2016). Recurrent hybridization and recent origin obscure phylogenetic relationships within the 'white-headed'gull (*Larus* sp.) complex. *Molecular Phylogenetics and Evolution* 103, 41–54. doi:10.1016/j.ympev.2016.06.008

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi:10.1093/bioinformatics/btu033

Stec, D., Krzywański, Ł., Zawierucha, K., and Michalczyk, Ł. (2020). Untangling systematics of the *Paramacrobiotus areolatus* species complex by an integrative redescription of the nominal species for the group, with multilocus phylogeny and species delineation in the genus *Paramacrobiotus. Zoological Journal of the Linnean Society* 188(3), 694–716. doi:10.1093/zoolinnean/zlz163

Strand, M., and Sundberg, P. (2011). A DNA-based description of a new nemertean (phylum Nemertea) species. *Marine Biology Research* 7, 63–70. doi:10.1080/17451001003713563

Strand, M., Norenburg, J., Alfaya, J. E., Ángel Fernández-Álvarez, F., Andersson, H. S., Andrade, S. C. S., Bartolomaeus, T., Beckers, P., Bigatti, G., Cherneva, I., Chernyshev, A., Chung, B. M., von Döhren, J., Giribet, G., Gonzalez-Cueto, J., Herrera-Bachiller, A., Hiebert, T., Hookabe, N., Junoy, J., Kajihara, H., Krämer, D., Kvist, S., Magarlamov, T. Y., Maslakova, S., Mendes, C. B., Okazaki, R., Sagorny, C., Schwartz, M., Sun, S. C., Sundberg, P., Turbeville, J. M., and Xu, C. M. (2019). Nemertean taxonomy – implementing changes in the higher ranks, dismissing Anopla and Enopla. *Zoologica Scripta* 48, 118–119. doi:10.1111/zsc.12317

Strong, E. E., and Whelan, N. V. (2019). Assessing the diversity of western North American *Juga* (Semisulcospiridae, Gastropoda). *Molecular Phylogenetics and Evolution* 136, 87–103. doi:10.1016/j.ympev.2019.04.009

Sundberg, P., Vodoti, E. T., Zhou, H., and Strand, M. (2009*a*). Polymorphism hides cryptic species in *Oerstedia dorsalis* (Nemertea, Hoplonemertea). *Biological Journal of the Linnean Society. Linnean Society of London* 98, 556–567. doi:10.1111/j.1095-8312.2009.01310.x

Sundberg, P., Chernyshev, A. V., Kajihara, H., Kånneby, T., and Strand, M. (2009*b*). Character-matrix based descriptions of two new nemertean (Nemertea) species. *Zoological Journal of the Linnean Society* 157, 264–294. doi:10.1111/j.1096-3642.2008.00514.x

Sundberg, P., Andrade, S. C. S., Bartolomaeus, T., Beckers, P., von Döhren, J., Krämer, D., Gibson, R., Giribet, G., Herrera-Bachiller, A., Junoy, J., Kajihara, H., Kvist, S., Kånneby, T., Sun, S. C., Thiel, M., Turbeville, J. M., and Strand, M. (2016*a*). The future of nemertean taxonomy (phylum Nemertea) – a proposal. *Zoologica Scripta* 45, 579–582. doi:10.1111/zsc.12182

Sundberg, P., Kvist, S., and Strand, M. (2016*b*). Evaluating the utility of single-locus DNA barcoding for the identification of ribbon worms (Phylum Nemertea). *PLoS One* 11, e0155541. doi:10.1371/journal.pone.0155541

Toews, D. P. L., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* 21, 3907–3930. doi:10.1111/j.1365-294X.2012.05664.x

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. doi:10.1093/bioinformatics/btt499

Zhbannikov, I. Y., Hunter, S. S., Foster, J. A., and Settles, M. L. (2017). SeqyClean: a pipeline for high-throughput sequence data preprocessing. In 'Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics',August 2017,Boston, MA.) pp. 407–416. (Association for Computing Machinery: New York, NY .)

Supplemental files
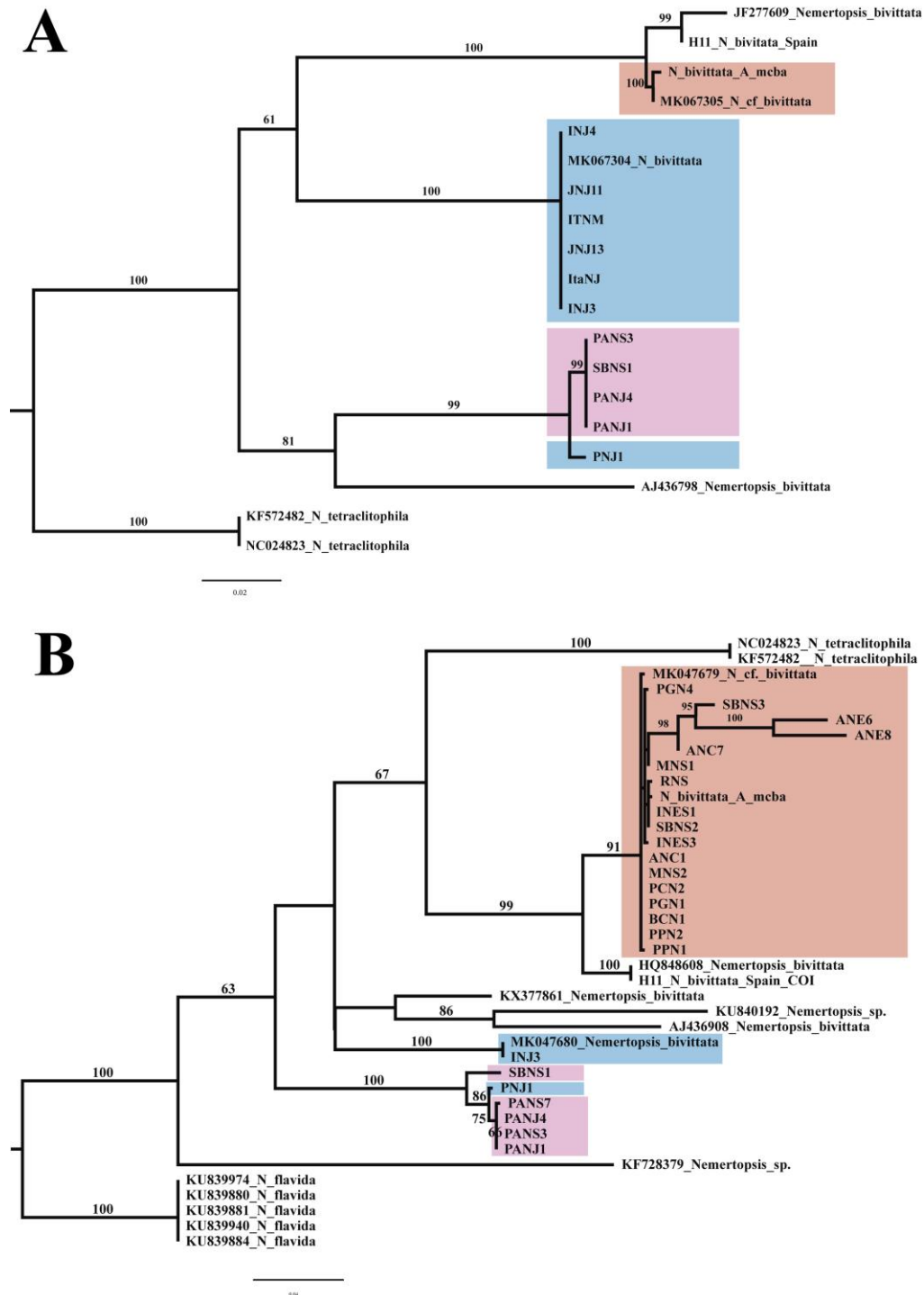
Figure S1. RAxML maximum likelihood results of *16S* rRNA (lnL = -1392.7393) and *COI* (lnL = -3199.6158) including all *Nemertopsis* sequences available in GenBank. Light orange shadow marks branch with specimens of *Nemertopsis berthalutzae* sp. nov.; light pink shadow marks branch with specimens of *Nemertopsis caete* sp. nov.; blue shadow marks branch with specimens of *Nemertopsis pamelaroeae*, sp. nov; A: results from 16rRNA sequences. B: results from *COI* sequences.
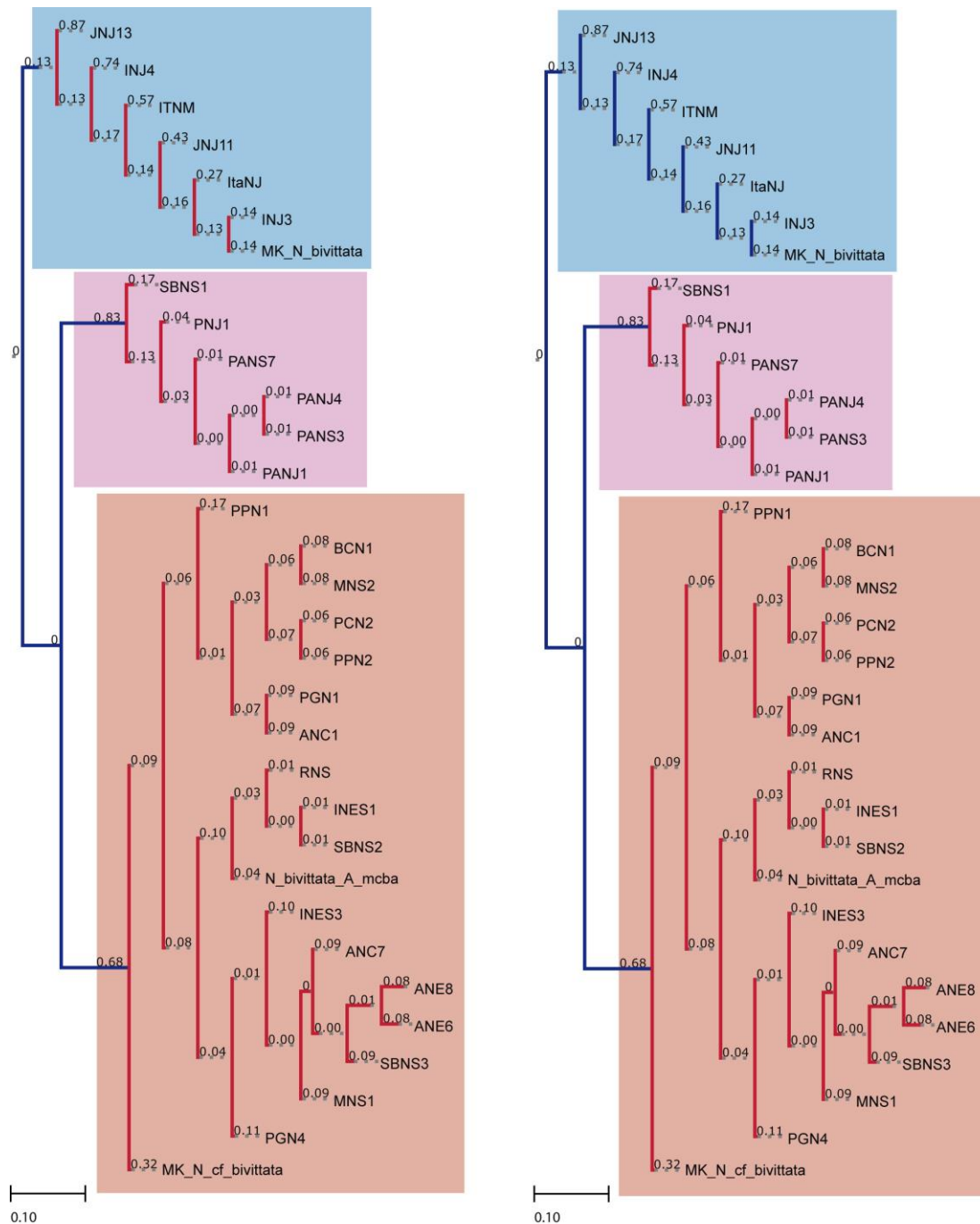
Figure S2. Species delimitation results from PTP for maximum likelihood (A) and Bayesian (B) solutions. Red branches indicates all terminals belong to the same species. Values on branches indicate probability of that branch comprises a single species. Light orange shadow marks branch with specimens of *Nemertopsis berthalutzae* sp. nov.; light pink shadow marks branch with specimens of *Nemertopsis caete* sp. nov.; blue shadow marks branch with specimens of *N. pamelaroeae* sp. nov.

Table S1. GenBank accession numbers for *COI* and *16S* rRNA sequences produced in this study *, numbers of topogenetypes

Sample name	Species name	Collection date	Location name	*COI* Accession	*16S* Accession
PANS3	*Nemertopsis caete* sp. nov.	Jan-2019	Brazil: Alagoas	MT433993	MT434999
PANJ4	*Nemertopsis caete* sp. nov.	Jan-2019	Brazil: Alagoas	MT433994*	MT434998*
PANJ1	*Nemertopsis caete* sp. nov.	Jan-2019	Brazil: Alagoas	MT433992	MT435000
SBNS1	*Nemertopsis caete* sp. nov.	Jul-2017	Brazil: Ceara	MT433996	MT434997
ItaNJ1	*Nemertopsis pamelaroeae* sp. nov.	Jun-2018	Brazil: Espirito Santo	Not sequenced	MT435001
PNJ2	*Nemertopsis pamelaroeae* sp. nov.	Jan-2017	Brazil: Ceara	Not sequenced	MT434996
INJIV	*Nemertopsis pamelaroeae* sp. nov.	Apr-2017	Brazil: Sao Paulo	Not sequenced	MT435002
INJIII	*Nemertopsis pamelaroeae* sp. nov.	Apr-2017	Brazil: Sao Paulo	MT433997*	MT435003
ANCVII	*Nemertopsis berthalutzae* sp. nov.	Jan-2017	Brazil: Sao Paulo	MT433988	Not sequenced
SBNSIII	*Nemertopsis berthalutzae* sp. nov.	Jul-2017	Brazil: Ceara	MT433989	Not sequenced
MNSII	*Nemertopsis berthalutzae* sp. nov.	May–2017	Brazil: Rio de Janeiro	MT433979	Not sequenced
SBNSII	*Nemertopsis berthalutzae* sp. nov.	Jul-2017	Brazil: Ceara	MT433984	Not sequenced
ANEVI	*Nemertopsis berthalutzae* sp. nov.	Jan-2017	Brazil: Sao Paulo	MT433991	Not sequenced
ANEVIII	*Nemertopsis berthalutzae* sp. nov.	Jan-2017	Brazil: Sao Paulo	MT433990	Not sequenced
MNSI	*Nemertopsis berthalutzae* sp. nov.	May–2017	Brazil: Rio de Janeiro	MT433987	Not sequenced
ANCI	*Nemertopsis berthalutzae* sp. nov.	Jan-2017	Brazil: Sao Paulo	MT433975	Not sequenced
JNJ11	*Nemertopsis*	Feb-2019	Brazil: Rio	Not	MT435004

	pamelaroeae sp. nov.		de Janeiro	sequenced	
JNJ13	*Nemertopsis pamelaroeae* sp. nov.	Feb-2019	Brazil: Rio de Janeiro	Not sequenced	MT435005*
ITN	*Nemertopsis pamelaroeae* sp. nov.	Jun-2018	Brazil: Rio de Janeiro	Not sequenced	MT435006
PGN1	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Sao Paulo	MT433976*	Not sequenced
PPN2	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Sao Paulo	MT433977	Not sequenced
PCN2	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Sao Paulo	MT433978	Not sequenced
BCN1	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Rio de Janeiro	MT433980	Not sequenced
PPN1	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Rio de Janeiro	MT433981	Not sequenced
INES3	*Nemertopsis berthalutzae* sp. nov.	Apr-2017	Brazil: Sao Paulo	MT433982	Not sequenced
RNS	*Nemertopsis berthalutzae* sp. nov.	May–2017	Brazil: Santa Catarina	MT433983	Not sequenced
INES1	*Nemertopsis berthalutzae* sp. nov.	Apr-2017	Brazil: Sao Paulo	MT433985	Not sequenced
PGN4	*Nemertopsis berthalutzae* sp. nov.	Feb-2019	Brazil: Sao Paulo	MT433986	Not sequenced

Table S2. Number of sequenced reads per specimen before and after filtering steps, in the final consensus and final number of loci

Abbreviations per specimen: Paracuru (PNJ); Sabiaguaba (SBNS); Prado (PANS and PANJ); Gamboa (GBN); Itaoca (ItaNS and ItaNJ); Itaipu (ITN); Praia das Goiabas (MNS); Prainha (PPN); Barra do Corumbê (BCN); Jabaquara (JNJ); Igararecê Marina (INJ and INES); Pontal da Cruz Pier (PCN); Araçá (ANC and ANE); Balneário dos Trabalhadores (PGN); Ribeirão da Ilha (RNS)

Specimen	Number of reads before filtering	Number of reads passed filters	Number of reads in cluster	Number of reads in clusters passed filter of depth	Number of reads in consensus	Number of loci in assembly
ANCI	3563793	3563224	18582	8255	7987	1362
ANCII	2539908	2539118	18475	7753	7586	1256
ANCIII	3117582	3117324	16553	6533	6361	1316
ANCIV	6353677	6352868	41306	12127	11765	1443
ANCV	6387024	6386416	89124	27612	27079	1146
ANCVI	8591471	8589146	23197	8228	7933	1148
ANCVII	1516035	1515828	15756	7304	7129	862
ANEI	2395800	2395644	24798	7453	7311	1349
ANEIII	5909122	5908220	73165	24747	24159	1248
ANEIV	3632801	3631951	46422	14799	14515	1253
ANEV	5583340	5582546	31954	14311	13977	1335
ANEVI	3639293	3638712	36118	16218	15901	1261
ANEVIII	7436830	7434559	45502	20095	19671	1266
GBN	542132	542104	16739	3897	3836	412
INJI	2132986	2132817	33083	11012	10866	688
INJII	4511961	4511697	76469	21996	21555	821
INJIII	3365552	3365367	51299	15462	15045	828
INJIV	7191822	7191424	44569	21315	20977	808
ItaNJ	556441	556396	23017	7059	6956	417
ItaNLE	92291	92283	11594	1459	1433	257
ItaNS	625322	625222	14721	3417	3338	687

100

JNJI	820942	820898	21206	4491	4458	610
JNJII	77642	77636	4560	1191	1175	254
JNJIII	4087749	4087456	37201	9648	9474	808
JNJIV	1604627	1604397	13472	4869	4805	687
JNJV	8988290	8987595	30136	9100	8946	772
JNJVI	2015856	2015759	25817	6128	6024	599
JNJVII	3846306	3846136	26243	6436	6314	759
JNJVIII	5786515	5786301	28652	8499	8347	806
MNSI	4987192	4986920	43196	12269	12007	1368
MNSII	1790799	1790548	39806	9220	9016	1244
PANJ1	571126	571125	55960	6183	6069	1860
PANJ2	1088968	1088964	36189	7241	7120	2450
PANJ3	1178137	1178126	42572	8353	8216	2422
PANJ4	1458866	1458860	41210	8913	8727	2412
PANJ5	957772	957765	26289	5861	5778	2279
PANJ7	390668	390662	35605	5959	5868	2117
PANJ8	1254896	1254886	45837	11151	10985	2462
PANJ9	2229862	2229846	78260	11236	11023	2476
PANS1	1365385	1365372	59411	9207	9119	2502
PANS2	1008767	1008753	60480	10621	10514	2505
PANS3	1558236	1558215	63218	12725	12466	2473
PANS4	3217998	3217966	64617	11038	10859	2389
PANS5	1626467	1626449	61984	8991	8849	2315
PANS6	159424	159423	13471	2581	2537	895
PANS7	1701075	1701065	69363	9892	9679	2313
PANS8	1120149	1120143	34245	8158	8069	2491
PANS9	1026296	1026287	37739	8235	8155	2497
PNJ1	820537	820467	15392	2901	2839	401

SBNS4	1830908	1830630	14574	4237	4105	738
SBNSI	7743688	7743231	82645	31668	30880	1361
SBNSII	3005087	3003822	15213	5688	5513	1389
SBNSIII	4444532	4443599	34573	11474	11211	1179
Total of reads	153449945	153432168	2011579	535216	524527	

NEX file

Resulting SNP nexus matrix after all filtering steps used in this study. Samples abbreviations same as in Table S2.

Chapter 3. Seascape genetics and connectivity in a polychaete worm: disentangling the roles of a biogeographic barrier and environmental

Cecili B. Mendes, Cinthya S. G. Santos, Thadeu Sobral-Souza, Arian Dielactaquiz Santos, Dalton Kei Sasaki, Danilo Augusto Silva, Marcelo Dottori, Sónia C. S. Andrade

ABSTRACT

Aim

Seascape genetics integrates physical oceanography, spatial statistics, bionomics and genetic diversity, among others, to quantify the landscape composition effect on to population connectivity. We present a study using SNPs to investigate the population structure of the annelid polychaete *Perinereis ponteni* Kinberg, 1865 (Phyllodocida: Nereididae), commonly found on the Brazilian coast, in an attempt to understand how the environmental heterogeneity and paleodistribution can shape population dynamics.

Location

Intertidal zone along the Brazilian coast from the northeast to the southeast coast, covering more than 3,000 km.

Methods

Genotyping-by-sequencing libraries were constructed, resulting in 16,119 SNPs. These SNPs were used to investigate the population structure and genetic diversity. The environmental variables used in the seascape tests performed in the LEA package were downloaded from Bio-Clim and Bio-Oracle platforms. Niche-based models, paleoclimatic simulations and Oceanic Circulation modeling were also applied.

Results

Our findings indicate high levels of connectivity between the populations, with most of the genetic variation within populations showing an isolation-by-distance pattern. Our results also detected a genetic boundary in the southeastern region of Brazil,

around Cabo Frio (Rio de Janeiro). This region is characterized by an upwelling that occurs mostly during the summer months. The paleodistribution and demographic genetic simulations denote a spatial refuge in the southeast during the Last Glacial Maximum (21 kya), with an expansion of northern and southern regions around 4,500 generations ago. We detected 1,133 SNPs with frequencies associated with eight environmental factors, most of them related to temperature.

Main conclusions

Even a species with such a high gene flow capability still responds to biogeographic barriers, highlighting the importance of biotic and abiotic factors shaping population connectivity. It is also worth noting the importance of temperature in shaping these populations, indicating that future climate change and ocean warming can have huge impacts on coastal communities.

Keywords: Genetic diversity, ocean currents, ecological niche modeling, rocky shores, Nereididae, *Perinereis ponteni*.

INTRODUCTION

For many years, marine populations have been thought to be panmictic, with high rates of gene flow (Palumbi, 1994; Caley *et al.*, 1996), because of the lack of explicit oceanic geographical barriers. In addition, some barriers to gene flow in marine systems are not absolute, being spatially or seasonally determined, resulting in partially isolated populations (Palumbi, 1994). In most marine species, the larval phase is the dominant dispersal stage, and the pelagic larval duration is directly linked to dispersal ability (e.g. Levin, 2006; Cowen and Sponaugle, 2009; Selkoe and Toonen, 2011). Consequently, species with planktonic larval stages are more likely to present better connected populations than species with no larvae. Nevertheless, many studies have shown that complex interactions between species-specific traits and environmental characteristics result in different patterns of connectivity that cannot be explained solely by larval dispersal through ocean currents (e.g., Bernardi *et al.*,

1993; Cowen and Sponaugle, 2009; Bernatchez *et al.*, 2018). Environmental conditions such as temperature and salinity have been previously suggested to be of adaptive importance in shaping gene flow in many marine species (Nanninga *et al.*, 2014; Saenz-Agudelo *et al.*, 2015; Sandoval-Castillo *et al.*, 2017; Bernatchez *et al.*, 2018; Singh *et al.*, 2018).

Seascape genetic studies, especially those using single nucleotide polymorphisms (SNPs), are highly sensitive in detecting patterns of population isolation and local adaptation, particularly because of the high number of markers analyzed. Therefore, SNPs can detect even discrete changes in gene flow (e.g. Saenz-Agudelo *et al.*, 2015; Selkoe *et al.*, 2016; Bernatchez *et al.*, 2018), allowing to identify regions that are possibly involved in the adaptive process, which provides important insights into a changing environment. Among invertebrates, most studies focus on species of commercial importance, mostly mollusks and crustaceans. Therefore, these patterns remain unknown for many invertebrate groups—such as polychaetes—that are crucial in marine communities. Polychaeta is a very diverse group of invertebrates under Annelida, comprising burial, planktonic, tube-dwelling, and reef-building species (Pamungkas *et al.*, 2019; Nunes *et al.*, 2021). They are distributed from tropical to polar regions and are among the most abundant and species-rich marine metazoans in benthic environments (Fauchald and Jumars, 1979; Rouse and Pleijel, 2001; Díaz-Castañeda and Reish, 2009; Jumars *et al.*, 2015). *Perinereis ponteni* is a Nereididae polychaete species with a broad spatial distribution in the Atlantic Ocean, from Mexico to Brazil. Although these animals are more easily found under *Brachidontes* sp. banks, they also inhabit algae, oysters, and barnacle beds, acting as important generalists in these communities. Despite its ecological importance, there are no detailed studies on their reproduction. The species is known to present epitoky with a

planktonic stage and might have a lecithotrophic larval phase like most Nereididae species do (Rouse, 2000; Bakken *et al*., 2018). Previous studies using mitochondrial markers to investigate the population structure of *P. ponteni* on the Brazilian coast found a high level of mixture, with high rates of gene flow between all populations (Paiva *et al.*, 2019).

In this study, we expected to find high levels of gene flow between the Brazilian populations of *P. ponteni*, as previously reported. Any disruption in this pattern is likely a result of environmental factors possibly acting in the larval dispersion and post-settlement survival rates as well as the past demographic scenarios. This is one of the first studies to evaluate the environmental role on population connectivity of an important invertebrate species in the Southwest Atlantic Ocean. Here, we coupled genomic scale sequencing with ecological association tests, oceanic circulation and niche-based models, and paleoclimate simulations to understand the present and historic gene flow patterns, and how these can be affected by the changing environment.

MATERIAL AND METHODS

Sampling

The specimens were collected from oyster and barnacle clusters in eight localities along the Brazilian coast (Table 1), through the years 2017–2019. The clusters were removed from exposed rocks and pillars with a spatula and transported to the laboratory in plastic bags. In the laboratory, clusters were submerged in seawater for a few hours before being taken apart; subsequently, all worms were carefully removed with forceps. The specimens were relaxed with pulverized menthol dissolved in seawater and identified under a stereomicroscope. Subsequently, worms

were fixed in 99% ethanol and stored at -20 °C until DNA extraction. Specimens were collected under permits issued by the Institute Chico Mendes (ICMBio), numbers 55701 and 67004.

DNA extraction, library preparation, and sequencing

Genomic DNA was extracted from 61 specimens of *Perinereis ponteni* using the protocol described by Doyle and Doyle (1987). The quality and concentration of the DNA were assessed by gel electrophoresis, spectrometry (NanoDrop Lite spectrophotometer; Thermo Fisher Scientific), and fluorometry (Qubit 3.0 Fluorometer; Invitrogen). Approximately 100 ng of the extracted DNA per sample were sent to EcoMol Consultoria e Projetos Ltda (Piracicaba, Brazil) where GBS libraries were prepared according to the protocol described by Elshire *et al.,* (2011) and modified by Nunes *et al.* (2017). Briefly, DNA was digested with *PstI* (De Donato *et al.*, 2013), and barcoding adapters were ligated to the fragmented DNA. The ligated fragments were amplified using PCR. After amplification, fragments were quantified using the KAPA library quantification kit (Roche, São Paulo, Brazil), and fragment size was checked using a Bioanalyzer (Agilent Technologies, São Paulo, Brazil). Finally, the libraries were sequenced in 100 bp single-end fragments using the Illumina Hiseq 2500 platform at the Centro de Genômica Funcional ESALQ-USP.

Table 1. *Perinereis ponteni* number of samples (n), observed heterozygosity (H_O), expected heterozygosity (H_E), and fixation coefficient (F_{IS}) values across all loci per sample. Geographic regions are given in brackets. *p value < 0.05; **p value < 0.01; ***p values < 0.001. Abbreviations are as in Table 1.

Location	n	H_O	H_E	F_{IS}
SBP (NE)	9	0.149	0.196	0.132***
PAP (NE)	5	0.150	0.180	0.139*
GBP (SE)	9	0.148	0.198	0.108**
ItaP (SE)	10	0.139	0.195	0.128***
PP (SE)	6	0.165	0.197	0.133***
JP (SE)	2	0.148	0.145	0.155

DuP (SE)	10	0.144	0.200	0.121***
AP (SE)	10	0.160	0.208	0.104**

SNPs prospection and data analysis

Sequenced fragments were first filtered using Seqyclean v1.10.09 (Zhbannikov *et al.*, 2017), removing sequences with an average Phred quality score ≤ 20, contaminants, and adapters (based on the UniVec database; https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/). After filtering, Ipyrad v0.7.30 (Eaton and Overcast, 2020) was used for demultiplexing, quality and size filtering, fragment clustering, and SNP prospection. Reads with more than five Ns or shorter than 35 bp were discarded. The minimum read depth was set to six for calling consensus sequences within the samples, and the maximum depth was set to 100,000. The clustering threshold was set to 90%. All loci sharing heterozygotic sites across more than eight individuals were filtered out to avoid paralogs and the maximum number of SNPs per locus was set to 30. A locus had to be present in at least 50% of the samples to be retained in the final dataset. All other parameters were maintained at default values. SNPs were first visualized as an occupancy matrix using the divergent option in the Matrix Condenser tool (de Medeiros and Farrell, 2018); samples with less than 60% occupancy were removed. The resulting variant call format (VCF) files were used as input to VCF tools v.0.1.16 (Danecek *et al.*, 2011), which removed sites with 50% or less of missing genotypes over all individuals, followed by the PLINK software (Purcell *et al.*, 2007), which filtered SNPs for missing data (geno < 0.45), rare alleles (MAF > 0.01), and linkage disequilibrium (indep-pairwise 50 5 0.5). The resulting files were converted to each specific program format using the PGDSpider v.2.1.1.5 (Lischer and Excoffier, 2012).

To characterize each population, molecular diversity indices (nucleotide diversity [θ_π] and nucleotide differences [θ_S]) were calculated in Arlequin (Excoffier and Lischer, 2010) using 1,000 simulations. Arlequin was also used to calculate pairwise $F_{ST,}$ global and population-specific F_{IS} (Weir and Cockerham, 1984), and analysis of molecular variance (AMOVA). The multi-loci estimates of observed (H_O) and expected (H_E) heterozygosity, and global F_{ST} (Weir and Cockerham, 1984) were calculated using the Adegenet package (Jombart, 2008) in R software (R Core Time 2013).

Structure software (Pritchard *et al.*, 2000) was used to analyze both the complete SNP dataset and outlier SNPs (see below), using values of K from one to 16 with default settings, except for lambda and alpha values. The alpha value was set to one in all runs, and the lambda value was first estimated in one run with all other values fixed. The estimated lambda value was fixed for 10 other runs. The results from the 10 iterations with fixed lambda values were compacted and analyzed in the R package Structure Harvester (Earl, 2012), where the likelihood of each K was compared, and the most suitable K was selected. The selected K was used to generate a bar plot showing the most likely ancestry of each individual in the R package PopHelper (Francis, 2017).

The whole dataset was also used in the Adegenet package as input for the Mantel test (Mantel, 1967) to identify possible isolation-by-distance (IBD). Each collection site was considered a different population for this test. Genetic distance was inferred using pairwise F_{ST} (Weir and Cockerham, 1984). The smallest distance between the two sites was calculated using Google Maps (Google Corporation) as a straight line connecting the two points. The Mantel test was performed with 1,000 simulations. The genetic transition between populations through cline-fitting models was

performed in Hzar (Derryberry *et al.*, 2014), an R package that allows flexibility in fitting cline models of varying complexity owing to its modular design. This package uses the allele frequencies of each SNP and attempts to correlate its frequency with the geographic distribution, creating a cline-fitting model for each allele. Only the 100 SNPs with higher contribution to principal component analysis (PCA) structuring were used (Fig. S1), because they had the most contrasting frequencies among the populations. Sixteen different models including a null model were tested by combining different tails and scaling values for each SNP (Table S1). Three independent runs for each model were conducted, with each run consisting of 20,000 iterations following an initial burn-in phase of 500 iterations. For each locus, the model with the lowest Akaike information criterion (AIC) was kept. Only the results where the null model was not accepted have been discussed.

To identify SNPs under selection, the SNP matrix was used as input in BayeScan v. 2.0 (Foll and Gaggiotti, 2008). The settings were as follows: 20 pilot runs of 50,000 iterations followed by 100,000 simulations with prior odds of 10 and 5% false discovery rate (FDR). Each location was considered to be a different population. Possible associations between environmental variables and SNP frequency were assessed with latent factor mixed models tests (LFMM) (Frichot *et al*., 2013; Rellstab *et al.*, 2015) with the R package LEA (Frichot and François, 2015), using both the number of clusters defined by Structure ±2 and the number of sampling locations ±2. The environmental variables from each geographic location were downloaded from two databases: WordClim (www.worldclim.org) and Bio-Oracle (https://bio-oracle.org/), and used to perform two separate PCAs, each PCA with variables from each database (Table S2). The four variables that better explained the species

distribution from each database were selected for association analyses with the LFMM function.

The outlier SNPs identified by each program were pooled and used as input for population structure analyses in Structure software. The same was done with the SNPs not identified by any of the programs, constituting the neutral dataset. These datasets were input for PCA using the Adegenet package. The outlier SNPs were also blasted against the nt and nr databases from NCBI. The identified proteins and non-coding RNAs were searched in the UniProt database (https://www.uniprot.org) to identify their functions.

The recent migration matrix and demographic scenarios were estimated in Fastsimcoal v.2.6 (Excoffier and Foll, 2011) based on the whole SNP matrix (neutral and outlier SNPs). The input used was the minor allele site frequency spectrum parameter (SFS) computed with Arlequin with 1,000 bootstraps. The population expansion in the three different demographic scenarios was estimated from the SFS file. In the population expansion scenarios, the three clusters identified in the Structure software were used as populations, and three models were tested. The first model estimated population expansion from the northeast (cluster 1) with posterior migration to southeastern populations (clusters 2 and 3, Fig. 1A). The second considered population expansion from cluster 2, with posterior northward migration to the northeastern populations and southward migration to the other southeastern populations. The third model estimated population expansion from cluster 3 (the southernmost cluster) with posterior northward migration to populations of clusters 2 and 1. For each model, 50 independent replicates, each including 40 estimation loops with 300,000 coalescence simulations, were performed. The probability of each

model given the observed data was determined based on both the maximum likelihood value and AIC. The recent migration rates were estimated using each locality as a population, with 10 independent replicates, each including 40 estimation loops with 60,000 coalescence simulations and assuming current migration between all pairs of populations. The recent migration pattern was also evaluated using the estimated effective migration surface (EEMS) program. This program uses the concept of effective migration to model the relationship between genetics and geography, using geo-referenced genetic samples (Petkova *et al.*, 2016). In this program, two chains with 9,000,000 interactions each were run, with a burn-in of 200,000 and thinning of 20,000 in both cases. The convergence graph and other graphic output files were generated using rEEMS plots (Petkova *et al.*, 2016) in R.

The paleodistribution of *P. ponteni* was based on niche-based models that use known occurrences and climate variables to predict the species distribution (Ferrier and Guisan, 2006; Stigall, 2012; Alvarado-Serrano and Knowles, 2014). To compile the known species occurrences, we used our sampling information and searched the Web of Science® database (Institute of Scientific Information, Thomson Scientific) using the following keywords and Boolean command combinations: Perinereis; AND ponteni; AND distribution* occurrence* record. As a result, 23 occurrence sites were compiled, encompassing the entire species distribution (Table S3). We conducted a factorial analysis to select the environmental variables (see more information in Cortez *et al.*, 2021). Different algorithms are currently able to predict the species distribution (Araujo *et al.*, 2019) and the combined use with ensemble forecasting is recommended. Here, we used four algorithms corresponding to presence-only methods: i) envelope score (bioclim), ii) Gower distance, and presence-background methods, iii) MaxEnt, and iv) support vector machine (SVM). We used 75% of the

presence to build the models, and 25 to test. We used the true skill statistic (TSS) value for model evaluation (Allouche *et al.*, 2006) and the maximum sensitivity and specificity as thresholds. All models were built to current climate conditions and projected to the Last Glacial Maximum (LGM, 21 kya) and Holocene (6 kya) climate conditions.

The population connection through larval dispersion was simulated in OpenDrift (1.7.1), an open-source particle tracking framework developed by Dagestad *et al*. (2017), that uses the Eulerian velocity and thermohaline fields from an Eulerian model and a fourth-order Runge-Kutta (Edwards *et al*., 2006) scheme to transport particles within the domain. The OpenDrift module Larvalfish have been here adapted to the *Perineris ponteni* larvae, where the functions to fish growth (Folkvord, 2005) and larvae vertical migration were applied to fit the module to the maximum length of the organism (~310 μm) and maximum depth of vertical migration (-2.5 m). The adapted module considers the drift composed of ocean currents, wind drift at the surface, vertical mixing, vertical migration, the hatch from egg to larvae and its continuous growth, and the terminal velocity according to Sundby (1993).

The OpenDrift simulations consisted of one year (from January 1st of 2019 to January 2nd of 2020), with daily releases of 1000 particles (i.e. eggs) inside a 10 m radius at Espírito Santo (41 ºW, 22 ºS); Ubatuba (45.1014 ºW, 23.5218 ºS) and São Sebastião (45.3687 ºW, 23.7406 ºS). The amount of 1000 particles per site was chosen because of its optimization between statistical significance and computational cost of processing and analysis. Each simulation had a time-step of 10800 seconds (=3h) and vertical mixing resolution of 60 s. After 40 days of drifting, the larvae were considered dead if not stranded to the shore. The current fields were provided by the

forecast system PreAMar (Costa *et al.*, 2020), which simulates conservative variables (temperature, pressure and salinity), currents and sea level (Blumberg *et al.*, 2010). The choice of the southwest coast of Brazil to simulate the larval dispersion is due to two main factors: (i) it is in this region that a gene flow barrier occurs and; (ii) the distance from and between the other two data locations is very large and, therefore, the currents would not allow a displacement long enough to proper establish a direct connection between those populations.

RESULTS

Prospection of SNPs and genomic diversity

Sequenced libraries resulted in 253,624,109 reads (BioProject PRJNA750868), with reads per sample ranging from 607,246 to 13,351,934. After several filtering and clustering steps, the average consensus read number was 36,603 (Table S4). From these, we analyzed 16,119 SNPs from 61 specimens, with the SNPs per sample ranging from 7,874 to 15,360, and a median of 14,791 SNPs per individual. The SNPs revealed that the observed heterozygosity across all loci was significantly lower than expected in most localities (Table 2). The fixation coefficient F_{IS} values across all loci were significant in seven out of eight localities, ranging from 0.104 to 0.157 (Table 2).

Table 2. Sampling sites, GPS coordinates, and the number of specimens collected. The abbreviation for each location is in parentheses. NE- Northeast, SE-Southeast.

Sampling location	State (Region)	Latitude	Longitude	Date of Collection	# Specimens
Sabiaguaba (SBP)	Ceará (NE)	-3.7687	-38.4358	Jul 2017	9
Prado (PAP)	Alagoas (NE)	-9.6769	-35.7515	Jan 2019	5
Gamboa (GBP)	Espírito Santo (SE)	-20.8887	-40.7655	Jun 2018	9
Itaoca (ItaP)	Espírito Santo	-20.9050	-40.7770	Jun 2018	10

	(SE)				
Prainha (PP)	Rio de Janeiro (SE)	-23.1482	-44.6955	Feb 2019	6
Jabaquara (JP)	Rio de Janeiro (SE)	-23.2112	-44.7141	Apr 2017 and Feb 2019	2
Praia Dura (DuP)	São Paulo (SE)	-23.4945	-45.1654	Jan 2018	10
Araçá (AP)	São Paulo (SE)	-23.8175	-45.4067	Jan 2017 and Feb 2019	10

Population structure with all SNPs

The Structure results from the matrices with the 16,119 SNPs apportioned the specimens of *P. ponteni* in three clusters ($\Delta K = 539.49$). The first encompassed all specimens from the northeast, and samples from Itaoca (SE) and Gamboa (SE). The second had the remaining specimens from SE, except four specimens from Prainha (SE) that are in the third cluster (Fig. 1A). The same pattern was observed in the PCA analyses, with the same three groups delimited (Fig. 1B). The results of pairwise F_{ST} showed significant values in most comparisons, ranging from 0.013 to 0.052. The non-significant F_{ST} values ranged from 0.0037 to 0.0396 in the six comparisons (Table S5). For AMOVA analyses, we defined each sampling site as a population and the two major clusters defined by Structure as the groups, with specimens from Prainha (SE) included in the group with other localities from Rio de Janeiro and São Paulo (SE). The results revealed most of the variation within individuals (85.35%), followed by the variation among individuals in each population (11.9%) (Table 3). Global values of F_{ST}, F_{IS}, and F_{SC} are presented in Table 3.

Table 3. AMOVA results and global fixation indexes with all loci considered. Statistically significant index values are in bold.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Global fixation indexes
Among groups	1	612.943	5.235	2.12	$F_{ST} = 0.022$
Among populations within groups	6	1749.544	1.549	0.63	$F_{IS} = $ **0.122**
Among individuals within populations	53	14257.3	29.330	11.90	$F_{SC} = $ **0.006**
Within individuals	61	12831	210.344	85.35	-

Cline analyses

Of the 100 SNPs that contributed most to genetic structuring, 52 were outlier SNPs and 48 were neutral SNPs. The null model had the lowest AIC values for all outlier SNPs and 33 neutral SNPs. Therefore, these SNPs cannot produce a cline-fitting graph. Fifteen neutral SNPs presented the lowest AIC for model 1 (no tail and fixed scaling), these clines are presented in Figure 2. The center slope of these clines is approximately 2,500 km away from Sabiaguaba (NE), which is around Prainha (SE) (Fig. S2).

All clines showed soft slopes, ranging between 0.00000118 and 0.0031, with SNP 11837 showing the sharpest cline (slope = 0.0031; width = 111.84). The mean ΔP of the cline tail values was relatively high ($\Delta P = 0.5119$), indicating an important differentiation. Nonetheless, no SNPs had a complete change in allele frequency ($\Delta P = 1$).
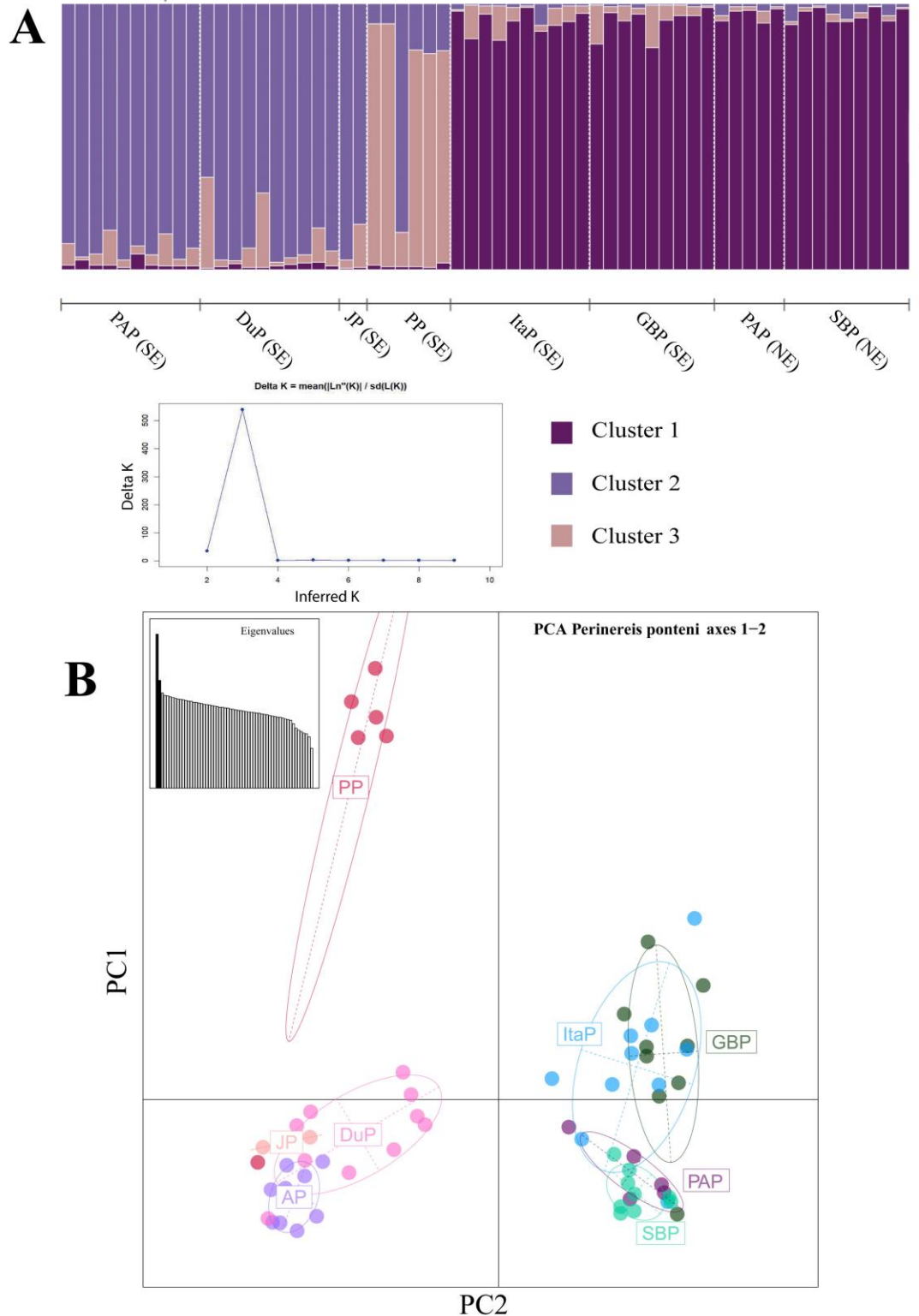
Figure 1. Clustering results for 16,119 SNPs from 61 *Perinereis ponteni* specimens. A. Bar chart showing Structure results for K = 3 and Structure Harvester results indicating K = 3 as the most probable K value. Each bar represents an individual and each color represents a different cluster. Horizontal lines delimit the sampling sites; B. PCA results, PC1 (eigenvalue = 1034.35) and PC2 (eigenvalue = 724.27). Each dot corresponds to one specimen. Different colors represent different sampling sites. Abbreviations are as in Table 1.

117

Demographic analyses and estimates of recent migration

The results from the demographic simulations in Fastsimcoal indicate that the population from Prainha (SE; Cluster 2 from Structure) was the most probable source for the other southeastern and northeastern populations (likelihood = -9787.992; Δ likelihood = 1570.977; AIC = 46421.0205; Fig. 2). The simulations also suggest that the split between the ancestral and the two current southeastern populations took place around 4,500 generations ago (estimated time of divergence = 4457). Migration rates show a stronger southward migrant flow (Fig. 3).

The recent migration estimates performed in the EEMS software indicate one probable barrier to migration around the Vitória-Trindade seamount chain (18°41'S 38°38'W), splitting the populations from northeast and southeast into two different groups, as indicated by different colors for each group (Fig. S3). In this case, populations from Espírito Santo (Gamboa and Itaoca) seem to exchange more migrants with other southeastern populations then with the northeaster ones. These results agree with the results of the Hzar cline test. The Mantel test showed a statistically significant correlation between genetic and geographic distance (R^2 = 0.393, p = 0.009), indicating IBD.

The larval dispersion simulation shows most of the larvae stay in the region where they were released, with only a few individuals transported to greater distances (Fig. 4). The individuals released in the vicinity of Itaoca (ES) seem to travel mostly southwards. They are transported only until right after the Cabo Frio region (around Saquarema, RJ), while the larvae that pass this area cannot reach any regions of the coast before the maximum time to metamorphosis (Fig 4A). The larvae released around Praia Dura (SP) and Araçá (SP), however, seem to mostly travel northwards,

reaching the coast of Rio de Janeiro in the vicinity of Cabo Frio (Fig. 4B and C). Nonetheless, the larvae released in this area (Araçá and Praia Dura) that travel southwards after the São Sebastião channel do not reach the coast in any area. This suggests that the Cabo Frio barrier is mainly determined by a hydrodynamic pattern (currents) rather than a thermodynamic one (temperature) associated with the local upwelling system
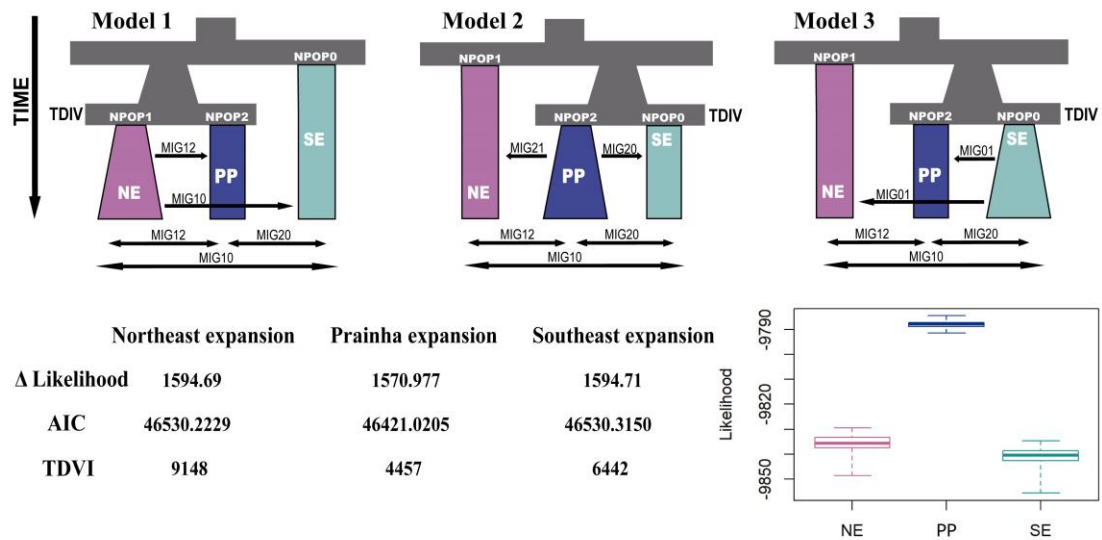
Figure 2. *Perinereis ponteni* demographic models simulated in Fastsimcoal (upper half) and their results (lower half). Models are presented backward in time, as represented by the black arrow on the left. Model 1 assumes a first split between the two northern populations, with subsequent migration southwards. Model 2 assumes a first split between the southeast populations, with a subsequent migration northwards and southwards. Model 3 assumes a first split between the two southeast populations, with a subsequent migration northward. The table below shows the comparison between the results from each model. Boxplot presents the individual likelihoods of each model. Each cluster is represented by the same color in all graphics. TDVI: time of divergence between two populations estimated by Fastsimcoal.

.
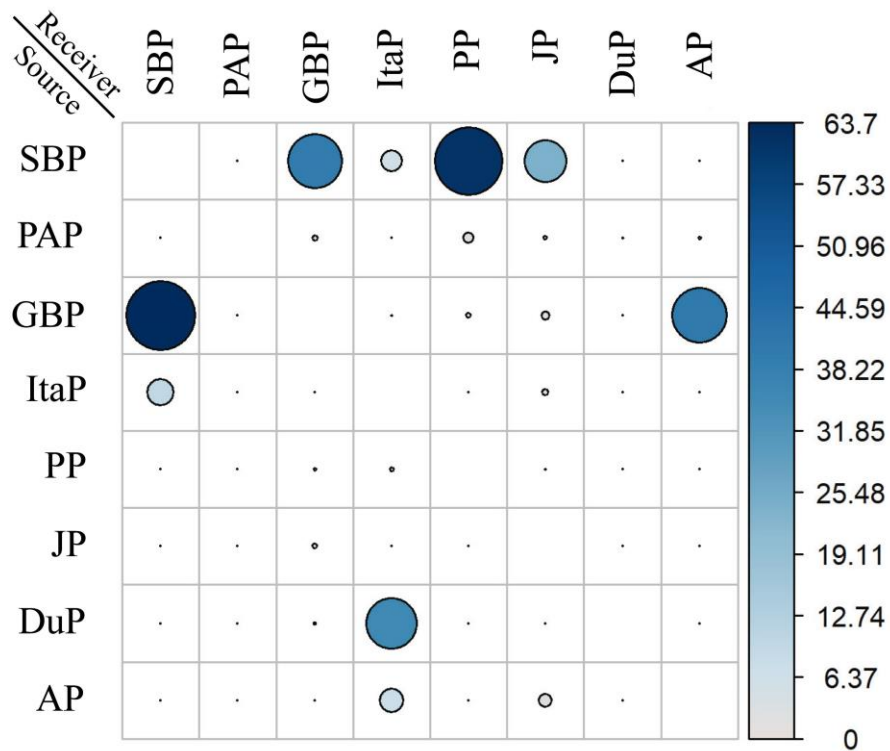
Figure 3. Migrant rates per generation from Fastsimcoal simulation with migration between all pairs of sites. Migration rate values are presented in the lateral bar. The circle sizes and colors represent the rates of migrants from the source locality (rows) toward the receiver locality (columns).

Outlier SNPs

BayeScan identified 107 outlier SNPs, all of which indicated positive selection (Fig. S4), while LFMM indicated the presence of 441 loci (1,133 SNPs) that had allele frequencies correlated to the eight environmental variables tested (Table 4, Fig. S5 and S6). The SNPs identified by the two methods showed no overlap. All the selected variables from WorldClim are related to temperature and have 570 SNPs with frequencies associated with their variance. The variables from Bio-Oracle are mostly nutrient-related, with only one exception, the pH level at the sea surface, which has the lowest number of SNPs associated with its variance. Most loci could not be identified in the BLAST searches and no loci related to pH or temperature seasonality were identified. The majority of the identified genes were involved in the normal cell cycle and muscular contraction (Table S6). We also found loci related to DNA repair

(*ZRANB3*), immunity (*PARP9*), epithelial growth (*CAPRIN2*), and germline differentiation (*TDRD7*), all correlated with chlorophyll A levels. In addition, the only outlier locus from BayeScan was identified as blasted against a gene with a function related to the development of the eyes (*FZD4*). The loci *TDR7* and *CAPRIN2* are also related to eye formation (Ciccia *et al.,* 2012; Mennella, 2015; Cipolla *et al*., 2016; Weston *et al*., 2021).

To evaluate any possible structuring related to the environmental variables, outlier SNPs were also used to investigate population structure. PCA results exhibited the same pattern seen for all loci, with tree clusters (not presented). Structure revealed only two clusters: the first was with specimens from SE locations up to Rio de Janeiro, and the second was with specimens from all other locations. The specimens from Prainha (the location with its cluster in analyses with all SNPs) have a mixed contribution from both clusters (Fig. S7). This is also the region where the cline pattern was detected for the neutral SNPs (Fig. S2). The Structure results using only the neutral SNPs also revealed three genetic clusters. In the neutral result, the first cluster comprehended all populations from the southern part of the Southeast, including Prainha, while the second cluster was more represented in the two populations from the northern part of the southeast (Gamboa and Itaoca), but these populations also show genetic contributions from the third cluster that is predominant amongst the northeast populations (Fig. S8).
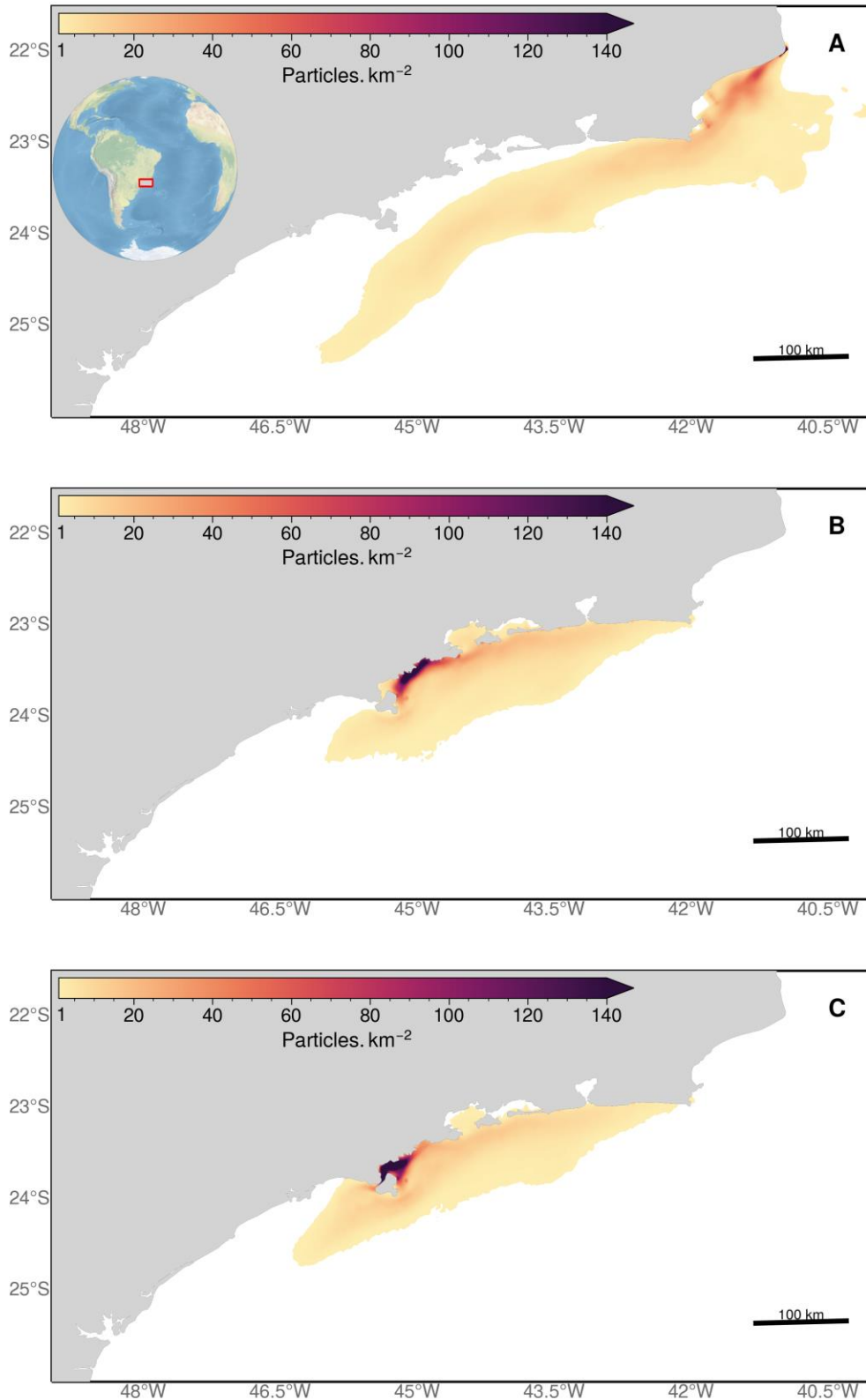
Figure 4. Population density considering 1000 daily seeded individuals in the vicinity of Itaoca - ES (A), Praia Dura - SP (B), and Araçá - SP (C). A hydrodynamic simulation PreAMar (Costa *et al*., 2020) provides the currents, salinity and temperature for the year of 2019. Each individual is viable for an interval of 50 days.

Table 4. Environmental variables selected for LFMM association. The number of SNPs associated with each variable is presented.

Source database	Variable #	Variable Name	# SNPs associated
Bio-Oracle	Variable 1	Chlorophyll A, maximum at the sea surface	147 SNPs
	Variable 2	Carbon phytoplankton biomass, mean at min depth	188 SNPs
	Variable 3	pH at the sea surface	163 SNPs
	Variable 4	Nitrate concentration, the range at min depth	108 SNPs
WorldClim	Variable 5	Temperature Seasonality (standard deviation ×100)	160 SNPs
	Variable 6	Minimum Temperature of Coldest Month	122 SNPs
	Variable 7	Mean Temperature of Wettest Quarter	118 SNPs
	Variable 8	Mean Temperature of Warmest Quarter	170 SNPs

Niche-based inferences

The simulations of the paleodistribution for *P. ponteni* suggest the presence of an ecological refuge around the southeast (~ 22° – 33□ S) region of Brazil during the Last Glacial Maximum (LGM, ~21 kya). This region that includes the Brazilian coast from Rio de Janeiro to the Rio Grande do Sul is the only one to present environmental conditions for species survival at that time (Fig. 5A). The suitability distribution was drastically modified in the mid-Holocene (~6 kya), when all the Southwest Atlantic coast, from the Caribbean up to the north of Rio Grande do Sul (South region), was highly suitable for species survival (Fig. 5B). The current spatial distribution simulation showed similar results, with all coasts highly suitable for *P. ponteni* (Fig. 5C). According to the suitability distribution, the southern coast of South America

and the Amazonas river estuary show lower values of suitability, indicating that these areas might act as permeable barriers.
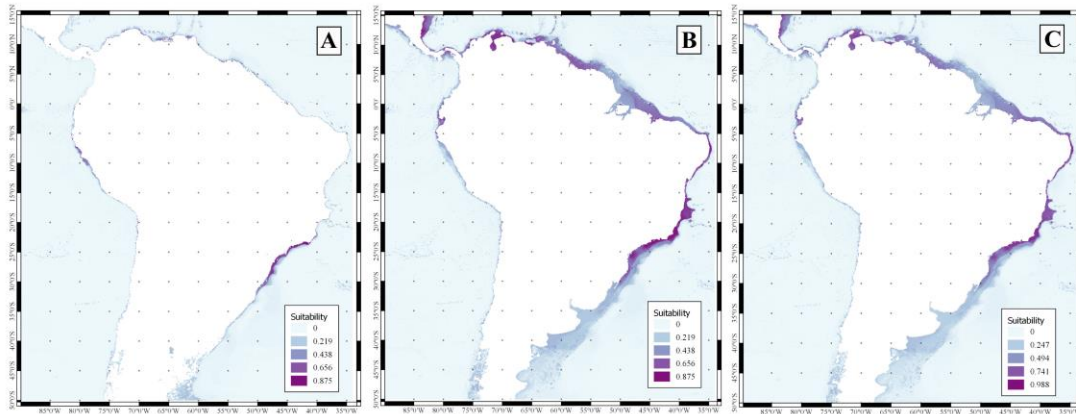
Figure 5. Niche modeling for paleodistribution of *Perinereis ponteni* from Last Glacial Maximum, ~ 21 kya (A); Mid-Holocene, ~ 6 kya (B); and present time, 0 kya (C). Different colors represent different values of suitability for the species as depicted in each image.

DISCUSSION

Here, we present the first attempt to assess the genomic diversity and possible factors influencing the population structure of a polychaete species along with larvae drift and niche modelling approaches. Our results indicate well-connected populations, but with an isolation-by-distance pattern, with less intense gene flow between the northeastern and southeastern regions. The findings also indicated a cline pattern of distribution for the 15 neutral SNPs. In the southeastern region, changes in allele distribution that characterize the cline occur in the same region where the gene flow between populations is weaker. The paleodistribution and demographic simulation results further suggest that this is the region where the species distribution was restricted during the mid-Holocene.

According to Rouse (2000), all Nereididae species have lecithotrophic larvae, and as such, they have a shortened pelagic larval duration (PLD) until settlement compared

124

to planktotrophic larvae (Todd and Doyle, 1981; Hoegh-Guldberg and Emlet, 1997). However, their PLD is still long enough for these larvae to travel thousands of kilometers connecting the Brazilian populations carried by ocean currents, since the adults do not usually travel great distances. The detected low and non-significant value of the global F_{ST} ($F_{ST} = 0.022$) is an indication of such connectivity. However, the IBD indicated by the significant results of the Mantel test and by the many significant pairwise F_{ST} values are an important evidence of the limits for dispersion of the *P. ponteni* larvae, probably due to factors other than the time of dispersion (Cowen and Sponaugle, 2009). The results from Structure and EEMS softwares further support this possibility, highlighting a boundary in the migration in the northern limit of the southeast region, either around Cabo Frio or around the Vitória-Trindade seamount chain, where something seems to act as a boundary for connectivity. The Structure, PCA, and cline analyses indicate that this boundary is located around the Cabo Frio region, while EEMS detects this boundary around the Vitória-Trindade seamount chain.

The Cabo Frio region is known to be a gene flow barrier or a sieve for many marine species, such as the crustaceans *Excirolana braziliensis* and *Litopenaeus schmitti* (Hurtado *et al*., 2016; Maggioni *et al.*, 2003), the coral *Mussismilia hispida* (Peluso *et al.*, 2018), the fish *Mugil liza* (Mai *et al.*, 2014), and even the dolphin *Pontoporia blainvillei* (Lázaro *et al.*, 2004). Mainly during the summer months, this region has strong winds coming from the continental area that push the warm coastal waters, giving rise to deeper cold waters rich in nutrients and forming a constant upwelling (Valentin *et al*., 1987; Martins *et al*., 2021). Therefore, the temperature and nutrient regimes in this area are quite different from those in subjacent zones, are thought to cause an alteration in the local community and acting as a semi-permeable boundary

for some species (Mai *et al.*, 2014; Peluso *et al.*, 2018; Volk *et al.*, 2020; Martins *et al.*, 2021). Our results, however, point to a hydrodynamic barrier preventing the larval exchange between adjacent areas (like Prainha and Gamboa), rather than an ecological one. A hydrodynamic barrier is the cause for isolation among coral species between the Great Barrier Reef and Lord Howe Island (Wood *et al.*, 2013; Keith *et al.*, 2015), and we believe this is also the cause for the pattern seeing here. The larval dispersion simulations show virtually no dispersion across the Cabo Frio region, but our demographic results show some migration among these sites. This sign can be an effect of past migration, or even migration events through stochastic events like storms. The Vitória-Trindade seamount is a chain of mostly submersed mounts that include the Abrolhos reef system. These mounts were exposed during the LGM, acting as a physical barrier to marine populations during this period (Thomas *et al.*, 2009; Martins *et al.,* 2021). With the sea rise during the Holocene (Lambeck *et al.*, 2014), the populations on different sides of this barrier were connected by ocean currents once again. Therefore, any genetic response to this barrier is an effect of past demographic dynamics. The populations of *P. ponteni* from Espírito Santo (Gamboa and Itaoca) and Rio de Janeiro (Prainha and Jabaquara) seem to have a stronger response to the Cabo Frio barrier, presenting significant pairwise values of F_{ST}, despite their proximity. This population structure, however, was not detected in a previous phylogeographic study using mitochondrial markers from Paiva *et al.* (2019). The authors analyzed the population structure of two *Perinereis ponteni* along the Brazilian coast, and the populations of *P. ponteni* presented panmixia with a few shared haplotypes distributed along the entire coast. Such disparity between genomic and mitochondrial data can be an effect of sex-biased inheritance associated with the different rates of mutation in the different markers (Toews and Brelsford, 2012;

Pazmiño *et al.*, 2017), as well as a result of the larger number of markers analyzed here.

The gene flow along the Brazilian coast resulting in succesive colonization events from different cohorts and generations might end in mixed genetically distinct subpopulations along the species distribution. This would explain the heterozygosity deficiency found, with statistically significant values of F_{IS}. Heterozygote deficiencies are commonly found in marine invertebrates and can be a result of high rates of inbreeding in small populations, natural selection, or the Wahlund effect (Castric *et al.*, 2002; Whitaker, 2004). Inbreeding is unlikely in *P. ponteni* since these animals produce epitokes with thousands of gametes (Rouse, 2000). Both natural selection and the Wahlund effect are plausible explanations for the heterozygosity deficiency uncovered here, with the Wahlund effect possibly playing the main role as the populations can be composed of several subpopulations.

Paleodistribution and origin of present populations

A previous study on mitochondrial genes found evidence of an important founder effect in *P. ponteni* populations along the Brazilian coast (Paiva *et al.*, 2019) that is congruent with our findings in the paleodistribution simulations. These simulations suggest that the environmental conditions necessary for *P. ponteni* survival were restricted to the southeast coast of Brazil during the LGM (~21kya) that acts as a refuge for the species. The demographic simulations in Fastsimcoal also confirm these results, indicating that the northern region of Rio de Janeiro (23□ S) is the most probable dispersion center for the species after the LGM. Fastsimcoal results show that the expansion occurred approximately 4,500 generations ago. Since the generation time of Nereididae varies considerably, from a couple of months to years,

and it is directly affected by the environmental temperature (Hardege, 1999), we cannot pinpoint the time of divergence between populations in years. However, our results of both the paleodistribution and demographic simulations are congruent, suggesting that *P. ponteni* distribution was restricted to the southeastern coast of Brazil during the LGM, expanding northward and southward during the mid-Holocene. The same pattern was found for other coastal species in Brazil (Ayres-Ostrock *et al.*, 2019; Cortez *et al.*, 2021), and given the great change in oceanic temperature and sea level from the LGM to the present (Tierney *et al.*, 2020), such restriction in the distribution is expected.

The climatic conditions in the mid-Holocene along the range of *P. ponteni* distribution were more similar to the present conditions (Steig, 1999). The ocean currents and ecological factors were more similar to the present ones as well (Gu *et al.*, 2018), and the species distribution in this period was very similar to the present. In addition, Fastsimcoal migration estimates show the same pattern, with migration southwards being orders of magnitude stronger than northwards. This pattern is likely a result of species dispersion through the Current of Brazil, a surface current that originates from the South Equatorial Current in northeast Brazil and flows southward along the Brazilian coast (Talley, 2011).

Signals of adaptive selection

Of the 16,119 SNPs analyzed, 1,240 were putatively under selection according to BayeScan and LEA, representing over 450 loci. Local adaptation is common among marine invertebrates with low motility in the adult phase, due to the stronger influence of local conditions (e.g. Leiva *et al.,* 2019; Sandoval-Castillo *et al.*, 2017; Gleason and Burton, 2016; de Wit and Palumbi, 2013).

The clustering analyses using only the outlier SNPs had a similar pattern to the analyses using the complete dataset and also using only the neutral SNPs (Fig. S7). This similar pattern may be a result of the geographic distribution of the environmental factors that affect this species, mainly the temperature and the availability of organic matter that have great latitudinal variance along the Brazilian coast (Grimm, 2011; Garreaud *et al.*, 2009). Our inability to identify many outlier loci is probably an effect of the lack of Nereididae annotated genomes that greatly affected the chance of finding hits in the online databases. However, the identified loci had hits to genes related to DNA repair (*ZRANB3*), immunity (*PARP9*), epithelial growth (*CAPRIN2* and *FZD4*), germline differentiation (*TDRD7*), and formation of the eyes (*FZD4*, *TDR7*, and *CAPRIN2*). *ZRANB3* prevents replication errors by stabilizing the replication fork (Ciccia *et al.*, 2012). This gene is also involved in DNA tolerance and repair. *PARP9* is an inactive enzyme from the *PARP* family that modulates the expression of interferon-stimulated genes, acting on the immune response to viral infections (Grimaldi *et al.*, 2016 and references therein). These genes may be related to the presence of anthropogenic contaminants in some areas where these animals were collected. They sometimes live in estuaries near fouling communities that are in many cases contaminated with domestic sewage and carry a variety of mutagenic substances and pathogenic microbiota.

Both *CAPRIN2* and *FZD4* are involved in regulating the canonical *Wnt* signaling pathway. This signaling controls tissue regeneration in the epithelium and retinal development (Kubo *et al.,* 2003; Goessling *et al.*, 2009). *TDR* is a family of proteins characterized by the Tudor domain repeats and is essential for germline differentiation in many animals, including *Drosophila melanogaster* and mice (Boswell and Mahowald, 1985; Chuma *et al.*, 2003; Hosokawa *et al.*, 2007). *TDR7* is

129

responsible for the formation of the cytoplasmic ribonucleoprotein complex together with *TDRD1/MTR-1* and seems to act on the correct positioning of germinal granules (Hosokawa *et al.*, 2007). However, the expression of *CAPRIN2* and *TRD7* is also related to ocular lens formation during eye development (Lachke *et al.*, 2011; Dash *et al.*, 2015), while *FZD4* expression has been recorded during rod photoreceptor differentiation in the adult mouse retina. In mice, the spatial and temporal patterns of expression of *FZD4* are consistent with their role in producing or stabilizing the rod photoreceptor (Hunter *et al.*, 2004). Since these three genes (*FZD*, *TDR7*, and *CAPRIN2*) can be related to different functions in different phases of development, and there is no data of their function in annelids, we cannot determine why they are presently under selection in *P. ponteni.*

This study sheds light on how temperature might influence connectivity among the Brazilian populations of *P. ponteni,* as indicated by the genetic boundary found in the upwelling region. Many of the outlier loci discovered here are directly or indirectly related to temperature. This is not only due to the disruption of local conditions, but also because the ocean currents are determined by the differences in ocean temperature and salinity, both of which can change dramatically in the coming years (Voosen, 2020; Zika *et al.*, 2018). This change in velocity and direction of the currents will most likely affect the connection between populations, affecting the gene flow in the species. However, not only temperature does affects species' survival or fitness, other environmental factors such as primary production may also have great effects on *Perinereis* fitness. This work reveals that polychaetes can also be affected at the genomic level by environmental changes and might be an interesting model for analysing the oceanic warming effect and its consequences on natural populations.

ACKNOWLEDGMENTS

REFERENCES

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, 43(6), 1223-1232.

Alvarado-Serrano, D. F., & Knowles, L. L. (2014). Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources*, 14(2), 233-248.

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., ... & Rahbek, C. (2019). Standards for distribution models in biodiversity assessments. *Science Advances*, 5(1), eaat4858.

Ayres-Ostrock, L. M., Valero, M., Mauger, S., Oliveira, M. C., Plastino, E. M., Guillemin, M. L., & Destombe, C. (2019). Dual influence of terrestrial and marine historical processes on the phylogeography of the Brazilian intertidal red alga *Gracilaria caudata. Journal of Phycology*, 55(5), 1096-1114.

Bakken, T., Glasby, C. J., Santos, C. S.G., & Wilson, R. S. R. S. (2018) Nereididae Blainville, 1818. In Handbook of Zoology - A Natural History of the Phyla of the Animal Kingdom - Annelida: Polychaetes. Editors, Westheide, W., Purschke, G. & Böggemann, M. De Gruytier.

Bernardi, G., Sordino, P., & Powers, D. A. (1993). Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. *Proceedings of the National Academy of Sciences*, 90(20), 9271-9274.

Bernatchez, S., Xuereb, A., Laporte, M., Benestan, L., Steeves, R., Laflamme, M., ... & Mallet, M. A. (2019). Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. *Evolutionary Applications*, 12(3), 587-609.

Blumberg, A., Georgas, N., Herrington, T., Bruno, M., 2010. How good is good? establishing performance goals for marine forecast systems. In: Ninth Conference on Coastal Atmospheric and Oceanic Prediction and Processes.

Boswell, R. E., & Mahowald, A. P. (1985). tudor, a gene required for assembly of the germ plasm in *Drosophila melanogaster*. *Cell*, 43(1), 97-104.

Caley, M. J., Carr, M. H., Hixon, M. A., Hughes, T. P., Jones, G. P., & Menge, B. A. (1996). Recruitment and the local dynamics of open marine populations. *Annual Review of Ecology and Systematics*, 27(1), 477-500.

Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. *Heredity*, 89(1), 27-35.

Ciccia, A., Nimonkar, A. V., Hu, Y., Hajdu, I., Achar, Y. J., Izhar, L., ... & Elledge, S. J. (2012). Polyubiquitinated PCNA recruits the ZRANB3 translocase to maintain genomic integrity after replication stress. *Molecular Cell*, 47(3), 396-409.

Cipolla, L., Maffia, A., Bertoletti, F., & Sabbioneda, S. (2016). The regulation of DNA damage tolerance by ubiquitin and ubiquitin-like modifiers. *Frontiers in Genetics*, 7, 105.

Chuma, S., Hiyoshi, M., Yamamoto, A., Hosokawa, M., Takamune, K., & Nakatsuji, N. (2003). Mouse Tudor Repeat-1 (MTR-1) is a novel component of chromatoid bodies/nuages in male germ cells and forms a complex with snRNPs. *Mechanisms of Development*, 120(9), 979-990.

Cortez, T., Amaral, R. V., Sobral-Souza, T., & Andrade, S. (2021). Genome-wide assessment elucidates connectivity and the evolutionary history of the highly dispersive marine invertebrate *Littoraria flava* (Littorinidae: Gastropoda). *Biological Journal of the Linnean Society*, blab055.

Costa, C.G.R., Leite, J.R.B., Castro, B.M., Blumberg, A. F., Georgas, N., Dottori, M., & Jordi, A. (2020). An operational forecasting system for physical processes in the Santos-São Vicente-Bertioga Estuarine System, Southeast Brazil. *Ocean Dynamics 70*(2), 257–271.

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, 1, 443-466.

Dagestad, K. F., Röhrs, J., Breivik, Ø, and Ådlandsvik, B. (2017). OpenDrift v1.0: a generic framework for trajectory modeling. *Geosci. Model Dev. Discuss*. 11, 1–28. doi: 10.5194/gmd-2017- 205

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158.

Dash, S., Dang, C. A., Beebe, D. C., & Lachke, S. A. (2015). Deficiency of the RNA binding protein caprin2 causes lens defects and features of peters anomaly. *Developmental Dynamics*, 244(10), 1313-1327.

de Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PloS One*, 8(5), e62137.

de Medeiros, B. A., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ*, 6, e5089.

de Wit, P., & Palumbi, S. R. (2013). Transcriptome-wide polymorphisms of red abalone (*H aliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, 22(11), 2884-2897.

Derryberry, E. P., Derryberry, G. E., Maley, J. M., & Brumfield, R. T. (2014). HZAR: hybrid zone analysis using an R software package. *Molecular Ecology Resources*, 14(3), 652-663.

Díaz-Castañeda, V., & Reish, D. (2009). Polychaetes in environmental studies. Annelids as model systems in the biological sciences. *Journal Wiley & Sons*, 205-227.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19(1), 11-15.

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.

Eaton, D. A., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, 36(8), 2592-2594.

Edwards, K. P., Werner, F. E., & Blanton, B. O. (2006). Comparison of Observed and Modeled Drifter Trajectories in Coastal Regions: An Improvement through Adjustments for Observed Drifter Slip and Errors in Wind Fields, *Journal of Atmospheric and Oceanic Technology*, *23*(11), 1614-1620.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, 6(5), e19379.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564-567.

Excoffier, L., & Foll, M. (2011). Fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27(9), 1332-1334.

Fauchald, K., & Jumars, P. A. (1979). The diet of worms: a study of polychaete feeding guilds. *Oceanography and Marine Biology Annual Review*, 17, 193-284.

Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43(3), 393-404.

Folkvord, A. 2005. Comparison of Size-at-Age of Larval Atlantic Cod (*Gadus Morhua*) from Different Populations Based on Size- and Temperature-Dependent Growth Models. *Canadian Journal of Fisheries and Aquatic Sciences*. *Journal Canadien Des Sciences Halieutiques et Aquatiques 62*(5): 1037-52.

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.

Francis, R. M. (2017). pophelper: a R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27-32.

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30(7), 1687-1699.

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925-929.

Garreaud, R. D., Vuille, M., Compagnucci, R., & Marengo, J. (2009). Present-day south american climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 281(3-4), 180-195.

Gleason, L. U., & Burton, R. S. (2016). Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*. *Molecular Ecology*, 25(15), 3557-3573.

Goessling, W., North, T. E., Loewer, S., Lord, A. M., Lee, S., Stoick-Cooper, C. L., ... & Zon, L. I. (2009). Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell*, 136(6), 1136-1147.

Grimaldi, G., Corda, D., & Catara, G. (2016). From toxins to mammalian enzymes: the diverse facets of mono-ADP-ribosylation. *Frontiers in Bioscience*, 20, 389-404.

Grimm, A. M. (2011). Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change. *Stochastic Environmental Research and Risk Assessment*, 25(4), 537-554.

Gu, F., Chiessi, C. M., Zonneveld, K. A., & Behling, H. (2018). Late Quaternary environmental dynamics inferred from marine sediment core GeoB6211-2 off southern Brazil. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 496, 48-61.

Hardege, J. D. (1999). Nereidid polychaetes as model organisms for marine chemical ecology. *Hydrobiologia*, 402, 145-161.

Hoegh-Guldberg, O., & Emlet, R. B. (1997). Energy use during the development of a lecithotrophic and a planktotrophic echinoid. *The Biological Bulletin*, 192(1), 27-40.

Hosokawa, M., Shoji, M., Kitamura, K., Tanaka, T., Noce, T., Chuma, S., & Nakatsuji, N. (2007). Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: domain composition, intracellular localization, and function in male germ cells in mice. *Developmental Biology*, 301(1), 38-52.

Hunter, D. D., Zhang, M., Ferguson, J. W., Koch, M., & Brunken, W. J. (2004). The extracellular matrix component WIF-1 is expressed during, and can modulate, retinal development. *Molecular and Cellular Neuroscience*, 27(4), 477-488.

Hurtado, L. A., Mateos, M., Mattos, G., Liu, S., Haye, P. A., & Paiva, P. C. (2016). Multiple transisthmian divergences, extensive cryptic diversity, occasional

long-distance dispersal, and biogeographic patterns in a marine coastal isopod with an amphi-American distribution. *Ecology and Evolution*, 6(21), 7794-7808.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403-1405.

Jumars, P. A., Dorgan, K. M., & Lindsay, S. M. (2015). Diet of worms emended: an update of polychaete feeding guilds. *Annual Review of Marine Science*, 7, 497-520.

Keith, S. A., Woolsey, E. S., Madin, J. S., Byrne, M., & Baird, A. H. (2015). Differential establishment potential of species predicts a shift in coral assemblage structure across a biogeographic barrier. *Ecography*, 38(12), 1225-1234.

Kubo, F., Takeichi, M., & Nakagawa, S. (2003). Wnt2b controls retinal cell differentiation at the ciliary marginal zone. *Development*, 130(3), 587-598.

Lachke, S. A., Alkuraya, F. S., Kneeland, S. C., Ohn, T., Aboukhalil, A., Howell, G. R., ... & Maas, R. L. (2011). Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. *Science*, 331(6024), 1571-1576.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., & Sambridge, M. (2014). Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proceedings of the National Academy of Sciences*, 111(43), 15296-15303.

Lázaro, M., Lessa, E. P., & Hamilton, H. (2004). Geographic genetic structure in the franciscana dolphin (*Pontoporia blainvillei*). *Marine Mammal Science*, 20(2), 201-214.

Leiva, C., Taboada, S., Kenny, N. J., Combosch, D., Giribet, G., Jombart, T., & Riesgo, A. (2019). Population substructure and signals of divergent adaptive selection despite admixture in the sponge *Dendrilla antarctica* from shallow waters surrounding the Antarctic Peninsula. *Molecular Ecology*, 28(13), 3151-3170.

Levin, L. A. (2006). Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*, 46(3), 282-297.

Lischer, H. E., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298-299.

Maggioni, R., Rogers, A. D., & Maclean, N. (2003). Population structure of *Litopenaeus schmitti* (Decapoda: Penaeidae) from the Brazilian coast identified using six polymorphic microsatellite loci. *Molecular Ecology*, 12(12), 3213-3217.

Mai, A. C., Mino, C. I., Marins, L. F., Monteiro-Neto, C., Miranda, L., Schwingel, P. R., ... & Vieira, J. P. (2014). Microsatellite variation and genetic structuring in *Mugil liza* (Teleostei: Mugilidae) populations from Argentina and Brazil. *Estuarine, Coastal and Shelf Science*, 149, 80-86.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2 Part 1), 209-220.

Martins, N., Macagnan, L., Cassano, V., & Gurgel, C. (2021). Barriers to gene flow along the Brazilian coast: a synthesis and data analysis. Authorea Preprints.

Nanninga, G. B., Saenz-Agudelo, P., Manica, A., & Berumen, M. L. (2014). Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, 23(3), 591-602.

Nunes, J. D. R. S., Liu, S., Pértille, F., Perazza, C. A., Villela, P. M. S., de Almeida-Val, V. M. F., ... & Coutinho, L. L. (2017). Large-scale SNP discovery and construction of a high-density genetic map of *Colossoma macropomum* through genotyping-by-sequencing. *Scientific Reports*, 7(1), 1-11.

Paiva, P. C., Mutaquilha, B. F., Coutinho, M. C. L., & Santos, C. S. (2019). Comparative phylogeography of two coastal species of *Perinereis* Kinberg, 1865 (Annelida, Polychaeta) in the South Atlantic. *Marine Biodiversity*, 49(3), 1537-1551.

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual review of ecology and systematics*, 25(1), 547-572.

Pamungkas, J., Glasby, C.J., Read, G.B. *et al.* (2019) Progress and perspectives in the discovery of polychaete worms (Annelida) of the world. *Helgoland Marine Research,* 73**(**4), 1-10

Pazmiño, D. A., Maes, G. E., Simpfendorfer, C. A., Salinas-de-León, P., & van Herwerden, L. (2017). Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). *Conservation Genetics*, 18(5), 1151-1163.

Peluso, L., Tascheri, V., Nunes, F. L., Castro, C. B., Pires, D. O., & Zilberberg, C. (2018). Contemporary and historical oceanographic processes explain genetic connectivity in a Southwestern Atlantic coral. *Scientific Reports*, 8(1), 1-12.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, 48(1), 94-100.

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

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559-575.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17), 4348-4370.

Rouse, G. W. (2000). Polychaetes have evolved feeding larvae numerous times. *Bulletin of Marine Science*, 67(1), 391-409.

Rouse G., Pleijel F.. (2001) Polychaetes. Oxford: Oxford University Press.

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.

Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W., & Beheregaray, L. B. (2017). Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. *Molecular Ecology*, 27(7), 1603-1620.

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., Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., ... & Toonen, R. J. (2016). A decade of seascape genetics: contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1-19.

Singh, S. P., Groeneveld, J. C., Hart-Davis, M. G., Backeberg, B. C., & Willows-Munro, S. (2018). Seascape genetics of the spiny lobster *Panulirus homarus* in the Western Indian Ocean: Understanding how oceanographic features shape the genetic structure of species with high larval dispersal potential. *Ecology and Evolution*, 8(23), 12221-12237.

Sundby, S. (1983). A one-dimensional model for the vertical distribution of pelagic fish eggs in the mixed layer Deep Sea Research (30) pp. 645-661.

Steig, E. J. (1999). Mid-Holocene climate change. *Science*, 286(5444), 1485-1487.

Stigall, A. L. (2012). Using ecological niche modelling to evaluate niche stability in deep time. *Journal of Biogeography*, 39(4), 772-781.

Talley, L. D. (2011). Descriptive physical oceanography: an introduction. Academic press.

Thomas, A. L., Henderson, G. M., Deschamps, P., Yokoyama, Y., Mason, A. J., Bard, E., ... & Camoin, G. (2009). Penultimate deglacial sea-level timing from uranium/thorium dating of Tahitian corals. *Science*, 324(5931), 1186-1189.

Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., & Poulsen, C. J. (2020). Glacial cooling and climate sensitivity revisited. *Nature*, 584(7822), 569-573.

Todd, C. D., & Doyle, R. W. (1981). Reproductive strategies of marine benthic invertebrates: a settlement-timing hypothesis. *Marine Ecology Progress Series*, 4(7), 5-8.

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

Valentin, J. L., Andre, D. L., & Jacob, S. A. (1987). Hydrobiology in the Cabo Frio (Brazil) upwelling: two-dimensional structure and variability during a wind cycle. *Continental Shelf Research*, 7(1), 77-88.

Volk, D. R., Konvalina, J. D., Floeter, S. R., Ferreira, C. E., & Hoffman, E. A. (2021). Going against the flow: Barriers to gene flow impact patterns of connectivity in cryptic coral reef gobies throughout the western Atlantic. *Journal of Biogeography*, 48(2), 427-439.

Voosen, P. (2020). Climate change spurs global speedup of ocean currents. *Science*, 367(6478), 612-613.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

Weston, R., Peeters, H., & Ahel, D. (2012). ZRANB3 is a structure-specific ATP-dependent endonuclease involved in replication stress response. *Genes & Development*, 26(14), 1558-1572.

Whitaker, K. (2004). Non-random mating and population genetic subdivision of two broadcasting corals at Ningaloo Reef, Western Australia. *Marine Biology*, 144(3), 593-603.

Wood, S., Paris, C. B., Ridgwell, A., & Hendy, E. J. (2014). Modelling dispersal and connectivity of broadcast spawning corals at the global scale. *Global Ecology and Biogeography*, 23(1), 1-11.

Zhbannikov, I. Y., Hunter, S. S., Foster, J. A., & Settles, M. L. (2017). SeqyClean: a pipeline for high-throughput sequence data preprocessing. In Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics (pp. 407-416).

Zika, J. D., Skliris, N., Blaker, A. T., Marsh, R., Nurser, A. G., & Josey, S. A. (2018). Improved estimates of water cycle change from ocean salinity: the key role of ocean warming. *Environmental Research Letters*, 13(7), 074036.

SUPPLEMENTAL FILES
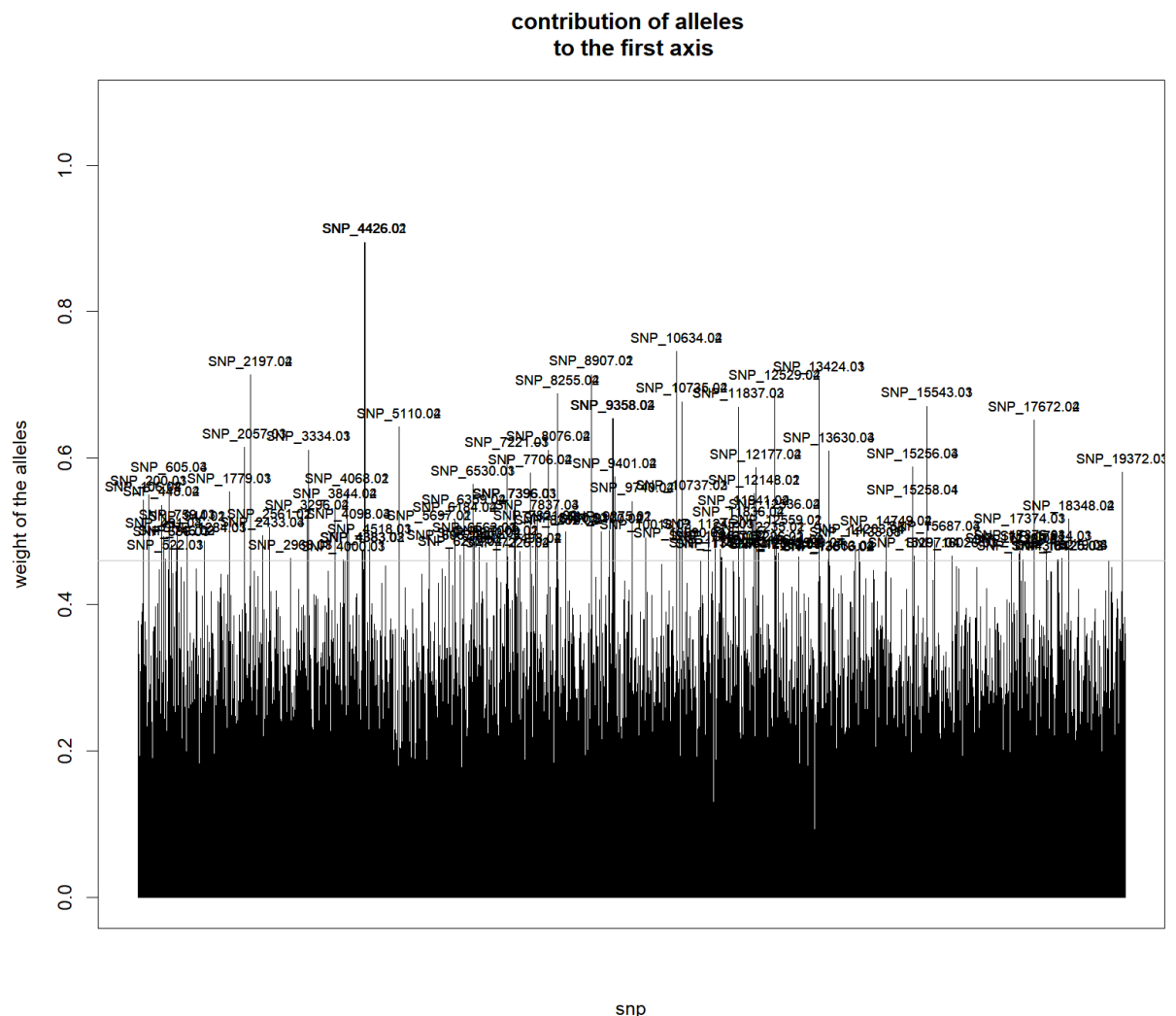
Figure S1. Loading plot showing the contribution of each SNP allele in the PCA of Perinereis ponteni. Each horizontal bar represents one allele of each SNP. The cut-off was set as 0.46 to select the 100 SNPs with higher contribution to PCA structure.
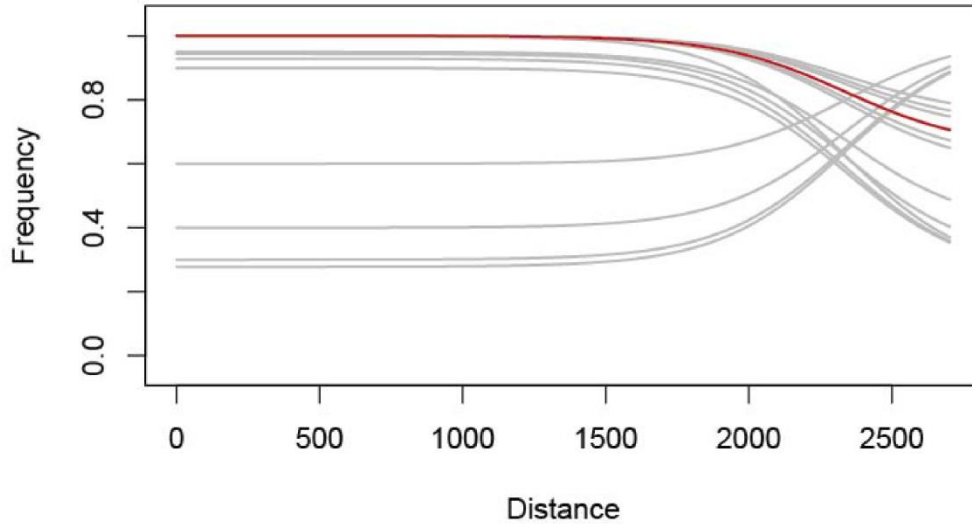
Figure S2. Hzar results for the best-fit model from the 16-model comparison in using the 15 neutral SNPs allele frequencies for *Perinereis ponteni*. In the plot of the maximum-likelihood cline, red dashed lines mark the sharpest slope cline
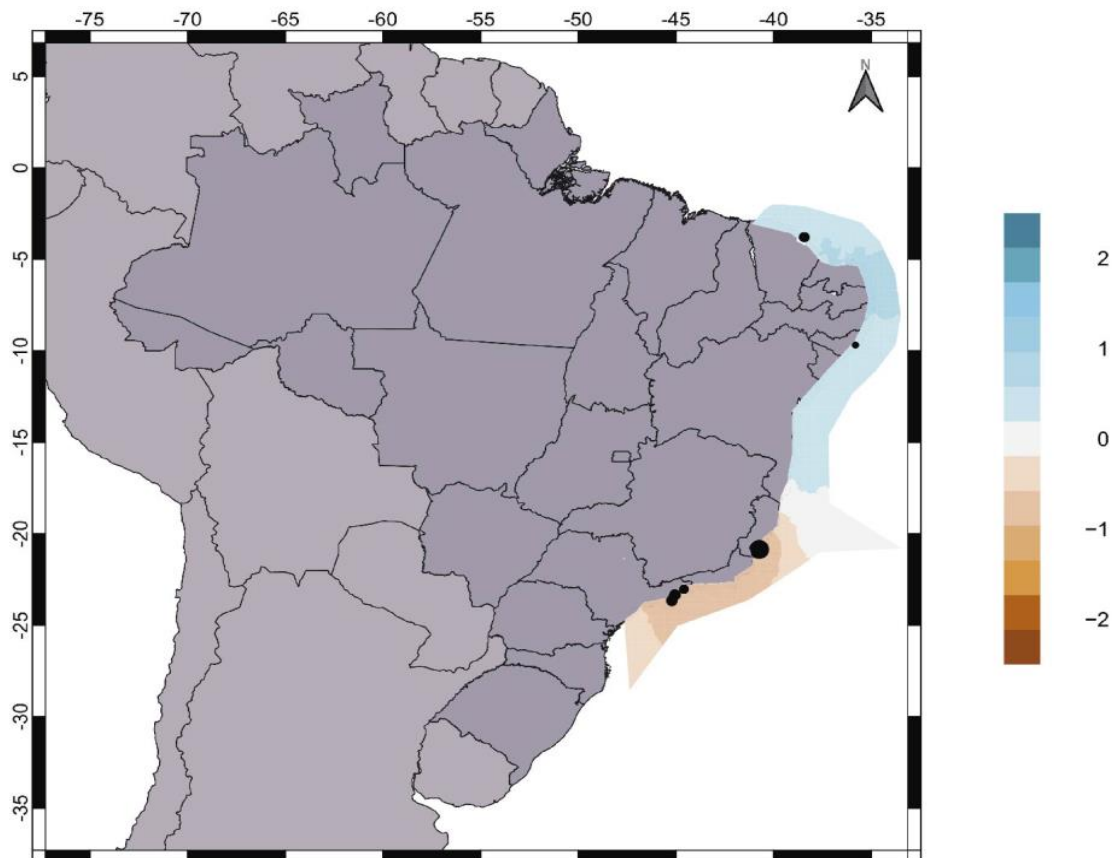
Figure S3. Estimated effective migration surface of *Perinereis ponteni*, on the log10 scale, after mean centering. Colors represent relative rates of migration, ranging from low (brown) to high (dark blue). Note that the two different color clusters separate the same two main clusters from Structure results, with NE partially isolated from the SE populations.

143

Figure S4. BayeScan plot of 16,199 SNPs from 61 specimens of Perinereis ponteni. F_{ST} is plotted against the log10 of the False Discovery Rate (FDR). Vertical line shows the critical FDR used for outlier identification (FDR = 0.1)

.

Figure S5. Manhattan plots of LFMM results using the four Bio-Oracle variables described in table 5 and 16199 SNPs from 61 specimens of Perinereis ponteni. Red dots represent outlier SNPs.

Figure S6. Manhattan plots of LFMM results using the four World Clim variables described in table 5 and 16199 SNPs from 61 specimens of Perinereis ponteni. Red dots represent outlier SNPs.

Figure S7. Structure Harvester results indicating K = 2 as the most probable K value (right) and a bar chart showing Structure results for K = 2 (left) for the outlier SNPs. Each bar represents an individual of *Perinereis ponteni* and each color represents a different ancestry. Horizontal lines delimit the sampling sites.
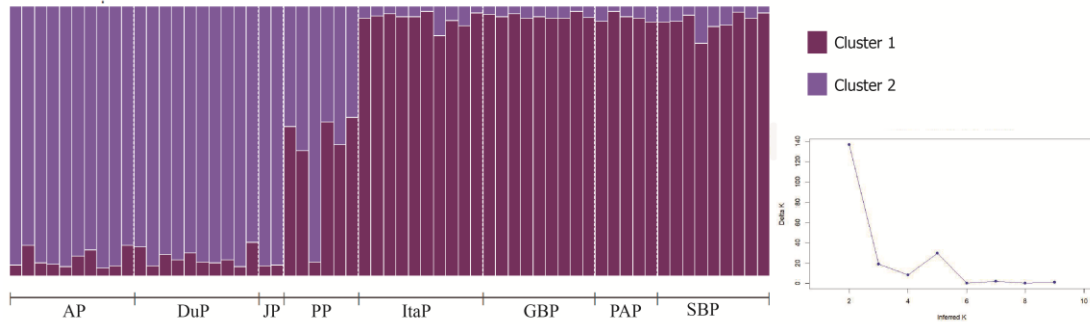
Figura S8. Structure Harvester results indicating K = 3 as the most probable K value (right) and bar chart showing Structure results for K = 3 (left) for the 14,879 neutral SNPs. Each bar represents an individual of *Perinereis ponteni* and each color represents a different ancestry. Horizontal lines delimit the sampling sites

Table S1. Initial parameter estimates for cline model construction on Hzar package. Types of scaling - fixed: model with minimum and maximum values fixed to the minimum and maximum observed mean values of data; none: model with fixed minimum value 0 and maximum value 1; free: model with the minimum and maximum value as free parameters. Types of tails – none: model with no exponential tails; right: A model with just one exponential tail on the right; left: model with just one exponential tail on the left; mirror: model with two exponential tails mirrored about the cline center; both: model with two tails with independent parameters

Model Name	Scaling	Tails
Model 1	fixed	none
Model 2	fixed	right
Model 3	fixed	left
Model 4	fixed	mirror
Model 5	fixed	both
Model 6	none	none
Model 7	none	right
Model 8	none	left
Model 9	none	mirror
Model 10	none	both
Model 11	free	none
Model 12	free	right
Model 13	free	left
Model 14	free	mirror
Model 15	free	both
Null model		

Table S2. Environmental variables used in Principal Component Analyses of *Perinereis ponteni* distribution from Bio-Oracle and World Clim databases.

Bio-Oracle Environmental Variables	
Layer Code	**Variable Name**
BO_chlomax	Chlorophyll A (maximum)
BO_chlomean	Chlorophyll A (mean)
BO_chlomin	Chlorophyll A (minimum)
BO_chlorange	Chlorophyll A (range)
BO_nitrate	Nitrate
BO_ph	pH
BO_sstrange	Sea surface temperature (range)
BO2_carbonphytomax_bdmin	Carbon phytoplankton biomass (maximum at min depth)
BO2_carbonphytomean_bdmin	Carbon phytoplankton biomass (mean at mean depth)
BO2_carbonphytorange_bdmin	Carbon phytoplankton biomass (range at min depth)
BO2_chlomin_bdmin	Chlorophyll concentration (minimum at min depth)
BO2_chlomax_bdmin	Chlorophyll concentration (maximum at min depth)
BO2_chlorange_bdmin	Chlorophyll concentration (range at min depth)
BO2_dissoxltmax_bdmin	Dissolved oxygen concentration (longterm max at min depth)
BO2_dissoxltmin_bdmin	Dissolved oxygen concentration (longterm min at min depth)
BO2_dissoxmax_bdmin	Dissolved oxygen concentration (maximum at min depth)
BO2_dissoxmean_bdmin	Dissolved oxygen concentration (mean at min depth)
BO2_dissoxmin_bdmin	Dissolved oxygen concentration (minimum at min depth)
BO2_dissoxrange_bdmin	Dissolved oxygen concentration (range at min depth)
BO2_nitratemax_bdmin	Nitrate concentration (maximum at min depth)
BO2_nitraterange_bdmin	Nitrate concentration (range at min depth)
BO2_salinitymean_ss	Sea surface salinity (mean)
BO2_temprange_bdmin	Sea water temperature (range at min depth)
BO2_temprange_ss	BO2_temprange_ss
World Clim Biological variables	
Layer Code	**Environmental Variable**
bio1	Annual Mean Temperature
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
bio3	Isothermality (BIO2/BIO7) ($\times 100$)
bio4	Temperature Seasonality (standard deviation $\times 100$)
bio5	Max Temperature of Warmest Month
bio6	Min Temperature of Coldest Month
bio7	Temperature Annual Range (BIO5-BIO6)

bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter

Table S3. Occurrence sites for *Perinereis ponteni* as available in publications from Web of Science® database.

Sampling location, state	Latitude	Longitude	Source of record
Paracuru, Ceará	-3.399704	-39.013765	Sousa, 2006 - Dissertation
Salinópolis, Pará	-0.697331	-47.370125	Chagas et al., 2018
Fortaleza, Ceará	-3.768708	-38.43586	Present study
Barra de Camaratuba, Paraíba	-6.603344	-34.964503	Paiva et al., 2019
Baía da Traição, Paraíba	-6.688583	-34.930508	Paiva et al., 2019
Tabatinga, Paraíba	-7.355767	-34.799122	Paiva et al., 2019
Tambaba, Paraíba	-7.366747	-34.797472	Paiva et al., 2019
Pina, Pernambuco	-8.089806	-34.879558	Paiva et al., 2019
Maceio, Alagoas	-9.676867	-35.751568	Present study
São Francisco do Conde, Bahia	-12.679969	-38.708236	Paiva et al., 2019
Ilha do Boi, Espírito Santo	-20.688281	-34.591222	Paiva et al., 2019
Gamboa, Espiríto Santo	-20.888769	-40.765536	Present study
Itaoca, Espírito Santo	-20.905027	-40.777093	Present study
Praia das Conchas, Rio de Janeiro	-22.871233	-41.980825	Paiva et al., 2019
Itaipu, Rio de Janeiro	-22.974086	-43.046942	Present study
Prainha, Paraty, Rio de Janeiro	-23.148219	-44.695563	Present study
Praia de Jabaquara, Paraty, Rio de Janeiro	-23.211277	-44.714141	Present study
Picinguaba, São Paulo	-23.357303	-44.865617	Paiva et al., 2019
Praia Dura, Ubatuba, São Paulo	-23.49265014	-45.1645267	Present study

Martin Sá, São Paulo	-23.625719	-45.375444	Paiva et al., 2019
Araçá, São Sebastião, São Paulo	-23.81756971	-45.40631776	Present study
Cananéia, São Paulo	-25.033333	-47.933333	Attolini et al., 1997
Ilha do Mel, Paraná	-25.5635	-48.318111	Ipucha et al., 2008

Table S4. Number of sequenced reads per specimen before and after filtering steps, in the final consensus and final number of loci. Abbreviations per specimen: Sabiaguaba (SBP); Prado Beach (PAP); Gamboa (GBP); Itaoca (ItaP); Prainha (PP); Jabaquara (JP); Dura Beach (DuP); Araçá (AP)

	# reads before filtering	# reads passed filters	# reads in cluster	# reads in clusters passed depth filter	# reads in consensus	# loci in assembly	#SNPs in final dataset
AP1	11344063	11343640	98327	53807	52933	6641	14946
AP10	4411843	4411691	72847	37609	36956	6532	15024
AP2	4838433	4838240	69089	35762	35169	6489	14615
AP3	3121747	3121622	62476	31502	30941	6214	14140
AP4	6706177	6705912	80664	43243	42464	6531	14900
AP5	5047669	5047511	71982	38829	38173	6480	14925
AP6	5013925	5013738	71769	38236	37587	6468	14723
AP7	7553081	7552705	88485	47411	46618	6583	15094
AP8	5346715	5346465	77570	41102	40415	6537	14865
AP9	3609083	3608896	67236	32738	32169	6306	14297
DuP1	2493796	2493610	53343	18433	18110	4496	10182
DuP10	8999225	8998679	76665	32600	32032	6591	15016
DuP2	2682188	2682021	57025	20315	19961	5003	11273
DuP3	5515381	5515070	67124	27862	27352	6276	14379
DuP4	2374811	2374682	46380	17480	17143	4159	9338
DuP5	1808705	1808610	44399	17397	17081	4306	9740
DuP6	2249692	2249572	46463	16264	15964	3819	8531
DuP7	4566762	4566532	60072	25032	24604	5920	13533
DuP8	11117537	11116887	79392	36239	35584	6630	15158
DuP9	13351934	13350968	96980	42035	41306	6677	15360
GBP1	1957418	1957324	57952	24335	23902	5841	13362

GBP10	2066722	2066621	61706	26228	25684	5859	13326
GBP2	2743630	2743429	70082	30498	29898	6181	13994
GBP3	3859788	3859620	71315	33486	32930	6574	15015
GBP4	2818528	2818378	64883	26848	26369	6131	14003
GBP5	5020979	5020772	75462	36564	35938	6618	15115
GBP7	4895045	4894826	84530	35552	34879	6605	15068
GBP8	4868990	4868756	76523	35742	35148	6606	15089
GBP9	4031784	4031586	75511	32946	32349	6543	14916
ItaP1	3478595	3478377	54242	20238	19852	4819	11044
ItaP10	5604169	5603916	85259	37568	36921	6638	15259
ItaP2	4042138	4041852	63230	27191	26721	6002	13763
ItaP3	6491658	6491234	72145	31118	30582	6484	14778
ItaP4	9028010	9027491	80470	33154	32591	6601	15115
ItaP5	5161054	5160695	66134	26931	26447	6108	13971
ItaP6	6422265	6421904	77311	30069	29501	6391	14427
ItaP7	1669738	1669441	53358	17476	17192	3586	7874
ItaP8	4192290	4192024	65941	26404	25927	5755	13119
ItaP9	3042443	3042322	68444	30134	29617	6376	14497
JP1	888534	888528	61564	26427	25990	5829	13332
JP2	4718295	4718270	138957	62536	61565	6619	15264
PAP1	2411668	2411649	92454	42020	41310	6543	14936
PAP2	1743375	1743357	94576	38992	38298	6515	14966
PAP3	607246	607233	50877	18452	18151	4443	9986
PAP4	4298338	4298300	135626	59679	58714	6586	15173
PAP5	5576002	5575957	144384	68738	67515	6583	15185
PP2	4359988	4359970	135515	58316	57315	6612	15217
PP3	4251463	4251447	141094	63102	62117	6605	15148
PP4	5667036	5666987	148756	67484	66431	6646	15339
PP5	7368488	7368457	173591	72680	71421	6618	15231
PP6	3030788	3030772	110525	53384	52563	6482	14916
PP7	3169027	3169005	114134	54199	53311	6568	15154
SBP1	6188686	6188410	94352	47253	46399	6538	14795

SBP10	7383054	7382802	86139	46205	45462	6572	15158
SBP2	7391498	7391201	89744	47693	46877	6572	15111
SBP3	2486116	2485965	64524	30650	30140	5703	13003
SBP4	5329541	5329359	82313	41780	41058	6561	14791
SBP5	3773103	3772821	87090	39792	39075	6439	14561
SBP6	5638896	5638656	80249	43084	42334	6544	14791
SBP7	2300639	2300554	62115	27930	27466	5594	12783
SBP8	6161489	6161200	86296	44981	44291	6503	14626

Table S5. Pairwise F_{ST} (below diagonal) and p-values (above diagonal) between all sampling sites with all loci considered. Statistically significant values of F_{ST} are in bold.

	SBP (NE)	PAP (NE)	GBP (SE)	ItaP (SE)	PP (SE)	JP (SE)	DuP (SE)	AP (SE)
SBP (NE)	*	0.829	0	0	0	0.018	0	0
PAP (NE)	0.008	*	0.135	0.018	0	0.090	0	0
GBP (SE)	**0.015**	0.013	*	0.874	0	0.009	0	0
ItaP (SE)	**0.016**	**0.019**	0.004	*	0	0	0	0
PP (SE)	**0.045**	**0.042**	**0.035**	**0.040**	*	0.117	0	0
JP (SE)	**0.042**	0.046	**0.052**	**0.049**	0.040	*	0.126	0.730
DuP (SE)	**0.032**	**0.034**	**0.037**	**0.033**	**0.026**	0.004	*	0.775
AP (SE)	**0.032**	**0.032**	**0.038**	**0.031**	**0.028**	0.016	0.026	*

Table S6. Genbank hits of the outlier SNPs identified through Blast searches against NT and NR databases.

Source of outlier SNP	Blast Hit	Species	Accession code	e-value
BayeScan	Frizzled class receptor 4 (FZD4)	*Zonotrichia albicollis*	XM_005485786.1	4.00E-10
Chlorophyll A	ARAP1	*Ochotona princeps*	XM_004589916.1	0.13
Chlorophyll A	TRIADDRAFT_16003	*Trichoplax adhaerens*	XM_002111084.1	0.079
Chlorophyll A	ACP7	*Erinaceus europaeus*	XM_007524022.1	0.097
Chlorophyll A	LINC02022	*Homo sapiens*	NR_136189.1	0.13
Chlorophyll A	LOC103360279	*Stegastes partitus*	XM_008285974.1	0.012
Chlorophyll A	CAPRIN2	*Homo sapiens*	NG_029557.2	0.029
Chlorophyll	COL1A2	*Anguilla japonica*	AB623214.1	0.13

A				
Chlorophyll A	TDRD7	*Cuculus canorus*	XM_009561179.1	0.2
Chlorophyll A	ZRANB3-like	*Pundamilia nyererei*	XM_005728430.1	0.097
Chlorophyll A	lacZ	*Mus musculus*	JN961734.1	0.2
Chlorophyll A	PARP9	*Bos mutus*	XM_005893415.1	0.095
Chlorophyll A	28S ribosomal RNA	*Culex pipiens*	GU911333.1	0.15
Chlorophyll A	COLlA1	*Ctenopharyngodon idella*	HM363526.1	0.097
Carbon phytoplankton biomass	Human Calmodulin	*Homo sapiens*	M19311.1	6.00E-06
Minimum Temperature of Coldest Month	DMWD	*Orycteropus afer afer*	XM_007943060.1	2.00E-06
Minimum Temperature of Coldest Month	Argonaute-1	*Blattella germanica*	HF912424.1	4.00E-10
Minimum Temperature of Coldest Month	mycbp2	*Callorhinchus milii*	XM_007890773.1	3.00E-09

Chapter 4. Comparative seascape genetics among sexual and partly asexual nemertean species reveals major influence of environment to local adaptation

Cecili B. Mendes, Jon L. Norenburg & Sónia C.S. Andrade

ABSTRACT

The reproductive mode and development strategies are key factors to the connectivity of marine populations, especially the benthic ones. Species with sexual reproduction and longer pelagic larval duration (PLD; e.g. planktotrophic) tend to present better connected populations than species with short PLD (e.g. lecithotrophic). In addition, asexual species can show less connected populations, and higher heterozygosis due to the accumulation of mutations (i. e. the Meselson effect). However, in a challenging environment, local adaptation can overcome other evolutionary forces in shaping the genomic patterns. To test these theories, we sampled the Brazilian populations of three ribbon worm species with different reproductive modes, *Lineus sanguineus*, *Nemertopsis berthalutzae* and *Nemertopsis pamelaroeae*. These are all coastal species that live in fouling communities, like oyster and barnacle beds. We used SNPs to characterize the populations, and compare the allele frequencies to environmental variables like SST and precipitation from two databases: WorldClim and Bio-Clim. Our results show well connected populations for the two planktotrophic species (*N. berthalutzae* and *N. pamelaroeae*), and for the partially asexual one (*L. sanguineus*), which also shows signs of isolation by distance. Nonetheless, all three species show SNPs candidate to selection in response to similar environmental variables, despite their differences. This points the environment great influence in shaping populations, raising once more the concern about the effect of climate change among coastal species.

Keywords: ribbon worms; fissiparity; dispersion; larvae; population genetics.

INTRODUCTION

Among marine animals, the mainly factor influencing the population connectivity seems to be the reproduction and development mode. However, other characteristics like the adult motility, population size and geographic barriers can also affect the connectivity (Cowen et al., 2007; Hellberg et al., 2002; Liggins et al., 2019; Palumbi, 1994). In benthonic species, the development mode is even more important for the connectivity, since these species have little or no motility. Usually larvae drift in surface ocean currents without being able to change their horizontal direction. However, they can change their horizontal position to remain in favorable environments or move towards the coast (Cowen & Sponaugle, 2009; D'Aloia et al., 2015; Levin, 2006). Species with a planktotrophic larva will usually have less structured populations than species with lecithotrophic larva or direct development (Becker et al., 2007; Cowen & Sponaugle, 2009). In addition, some species can have asexual reproduction, and among those, the patterns of connectivity will also change. These species usually present less connected populations, with higher than expected heterozygosity due to the accumulation of mutations in absence of segregation, the so call Meselson effect (Ament-Velásquez et al., 2016; Ceplitis, 2003).

Seascape genomics is a field that takes into account how the environment matrix and historic processes affect the present dynamics and connectivity of marine populations (Selkoe et al., 2016; Singh et al., 2018). This approach uses genome wide markers, such as Single Nucleotide Polymorphisms (SNPs), to search the correlation between such processes and the bionomic features of a species. The use of SNPs instead of more conventional markers allows correlation between genomic and environmental data, offering the material to detect markers candidate to selection, due to the nature and higher number of markers analyzed (Davey & Blaxter, 2010;

Liggins et al., 2019; Narum et al., 2013). The correlation between the environmental and genomic data is especially important because environmental variability also influences connectivity, since the environmental gradient can affect larval survival and post-settlement fitness (Nanninga et al., 2014; Saenz-Agudelo et al., 2015). The knowledge about local adaptation is also important to understand the molecular basis of micro-evolution induced by clime, allowing predictions about future scenarios.

Nemerteans are mostly marine worms, characterized by having an eversible proboscis housed in a cavity called rhynchocoel. The phylum has about 1,300 know species (Gibson, 1995; Kajihara et al., 2008; Norenburg et al., 2017) separated in three classes, Palaeonemertea, Pilidiophora and Hoplonemertea (Andrade et al., 2012; Strand et al., 2019; Thollesson & Norenburg, 2003). Most known nemertean species live in shallow waters (Norenburg et al., 2017), using its proboscis to prey upon many different animal species. Usually, they evert the proboscis towards the prey and the toxins present in the mucus paralyze it (Thiel & Kruse, 2001).

The nemertean reproduction and development is not entirely known, but most species are dioecious with external fertilization, producing a planktonic larva (Maslakova & Hiebert, 2014; Thiel & Junoy, 2006). Most pilidiophorans have a planktotrophic larva, the pilidium (Bird et al., 2014). Palaeonemerteans and hoplonemerteans, however, have a planuliform larva that resembles the adults and, because of it, they were considered as direct developers. Nevertheless, population genetic assessments, and more recently direct observation, indicate that many palaeonemertean and hoplonemertean larvae are planktotrophic and can stay in the plankton for several weeks (Andrade et al., 2011; Maslakova & Hiebert, 2014; Mendes et al., 2018; von Dassow et al., in press). Other nemertean species are partially or totally asexual, reproducing by fissiparity. These species can regenerate

both the posterior and anterior body parts, and are more common among the pilidiophorans (Ament-Velásquez et al., 2016; Ikenaga et al., 2019; Zattara et al., 2019).

Among the fissiparous nemerteans, the pilidiophoran *Lineus sanguineus* is one of the widest spread species. They are present in both North and South hemispheres, in the Atlantic and Pacific Oceans (Ament-Velásquez et al., 2016) and usually live among incrusting communities, like mussel mats, oyster and barnacles beds. The species seem to feed mainly upon annelid polychaetes, specially phyllodocids and syllids (Caplins & Turbeville, 2011; Thiel & Kruse, 2001). There are some reports of gonadal development, but no sexual reproduction was ever recorded (Gontcharoff, 1950; Moretto & Brancato, 1997). Nonethless, in the study of Ament-Velásquez et al. (2016), the authors could not find genetic evidence of asexuality as playing a main role the *L. sanguineus* populations, and infer that this species might be partially asexual.

The hoplonemerteans *Nemertopsis berthalutzae* and *Nemertopsis pamelaroeae* are part of *Nemertopsis bivittata* species complex, and live in the same habitat as *L. sanguineus*. These species are only known for the Brazilian Argentinean Coasts so far (Mendes *et al.,* 2021). Their reproduction is not entirely known, however, close related species living in similar habitats, such as *Emplectonema viride*, produce a long lived planktotrophic larvae (von Dassow *et al.*, in press; Mendes *et al.*, in prep).

To understand how different reproduction and developmental strategies can influence the population structure in a challenging environment as the intertidal zone, we collected these three nemertean species along more 3,000km in the Brazilian Coast. Our initial hypothesis was that 1) the partially asexual species show more structured populations than species with planktonic development, the same being true

for signs of isolation-by-distance, and 2) temperature is the main driver for local adaptation for all species, despite mode of reproduction, being most of the SNPs putatively under selection related to this variable in some way.

MATERIALS AND METHODS

Sampling strategy

The specimens were collected among oyster and barnacles clusters in 12 localities along the Brazilian Coast (Table 1; Fig. 1), through the years of 2017-2019. The clusters were removed from exposed rocks and pillars with a spatula, and transported to the lab in plastic bags. In the lab, clusters were submerged in seawater for a few hours before taken apart and had all worms carefully removed with forceps or brush. The specimens were placed in a petri dish, relaxed in a mixture 1:1 of $MgCl_2$ (stock ~ isotonic to seawater) and seawater, measured, photographed and identified under macroscope. Subsequently, worms were fixed in 99% ethanol and kept at -20ºC until DNA extraction. Specimens were collected under permits issued by Institute Chico Mendes (ICMBio) numbers 55701 and 67004.

Table 1. Sampling sites, GPS coordinates and number of specimens collected of each species. Abbreviation for each location and for each geographical region between parentheses.

Sampling location	Species	GPS Coordinates	Date of Collection	Species collected Abbreviation - # Specimens
Paracuru (NE)	*N. pamelaroeae*	3.3997 S 39.0137 W	Jan 2017	PNJ - 1
Sabiaguaba (NE)	*N. berthalutzae*	3.7687 S 38.4358 W	Jul 2017	SBNS – 3
	L. sanguineus			SBL - 7
Gamboa	*N. pamelaroeae*	20.8887 S 40.7655 W	Jun 2018	GBN - 1
Itaoca	*N. berthalutzae*	20.9050 S 40.7770 W	Jun 2018	ItaNS – 2
	N. pamelaroeae			ItaNJ - 1
Boca da Barra	*L. sanguineus*	22.5319 S 41.9401 W	Jul 2018	ROL - 8

157

Itaipu	*L. sanguineus*	22.9741 S 43.0471 W	Dec 2017	ITL - 7
Boa Viagem	*L. sanguineus*	22.9090 S 43.1311 W	Dec 2017	BVL - 9
Praia das Goiabas	*N. berthalutzae*	23.0258 S 44.5144 W	Feb 2019 and May 2017	MNS - 2
Jabaquara	*N. pamelaroeae*	23.2112 S 44.7141 W	Apr 2017 and Feb 2019	JNJ - 8
Igararecê Marina	*N. pamelaroeae*	23.7682 S 45.4009 W	Apr 2018	INJ - 4
Araçá	*N. berthalutzae*	23.8175 S 45.4067 W	Jan 2017 and Feb 2019	ANE/ANC – 13
	L. sanguineus			ALS - 12
Praia do Costão	*L. sanguineus*	24.3371 S 47.0021 W	Mar 2018	PBL - 10

DNA extraction, library preparation and sequencing

Genomic DNA was extracted using the Doyle and Doyle (1987) protocol from 53 specimens of *L. sanguineus*, 19 specimens of *N. berthalutzae* and 16 specimens of *N. pamelaroeae*. About 100 ng of genomic DNA per sample was sent EcoMol Consultoria e Projetos Ltda (Brazil), where genotyping-by-sequencing (GBS) libraries were prepared according to the protocol of Elshire *et al.* (2011) and modification of Nunes et al. (2017). Briefly, DNA was digested with PstI (De Donato et al., 2013), adapters were ligated, and fragments amplified. After amplification, fragments were quantified using KAPA library quantification kit (Roche Sequencing Solutions) and fragment size checked using Bioanalyzer (Agilent Technologies). Finally, the libraries were sequenced in 100 bp single-end fragments using Illumina Hiseq 2500 platform, at the Centro de Genômica Funcional ESALQ-USP.
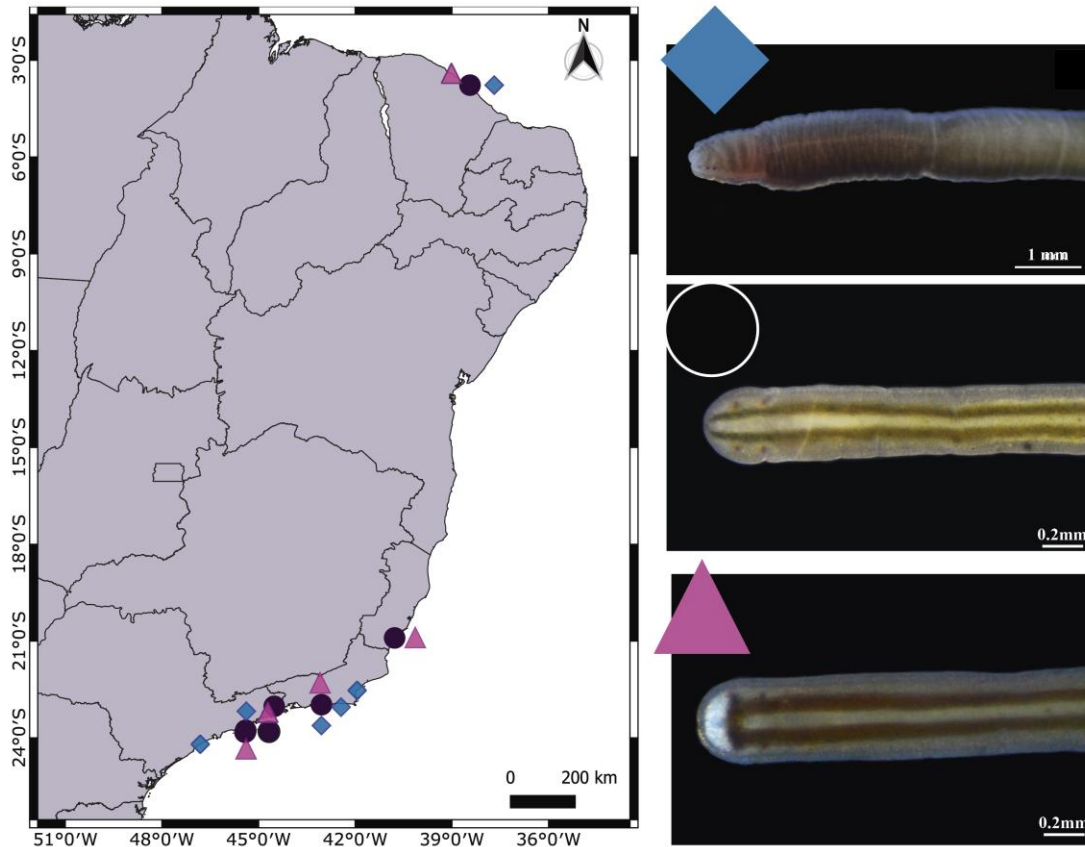
Figure 1. Map showing the sampling sites (left side) and representative of each species (right side). Blue diamonds represents *Lineus sanguineus*, black circles represents *Nemertopsis berthalutzae*, pink triangles represents *Nemertopsis pamelaroeae*.

SNP prospection, filtering and data analysis

The resulting raw reads were first filtered using Seqyclean v1.10.09 (Zhbannikov et al., 2017), removing sequences with average phred quality score ≤ 20, contaminants and adapters (based on the UniVec database; https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/). After filtering, SNP (Single Nucleotide Polymorphism) calling was performed using the de novo algorithm in Ipyrad v0.7.30 (Eaton & Overcast, 2020) as described in Mendes et al. (2021) for each species, separately. The final dataset was visualized as an occupancy matrix using the option divergent, in the Matrix Condenser tool (de Medeiros & Farrell, 2018) and samples with less than 60% occupancy were removed. The resulting VCF

159

(*Variant Call Format*) files were used as input to VCFtools v.0.1.16 (Danecek et al., 2011), where sites with 50% or more missing genotypes were removed, and then filtered in the PLINK software (Purcell et al., 2007) for missing data (geno < 0.45), rare alleles (MAF > 0.01) and linkage disequilibrium (indep-pairwise 50 5 0.5). The resulting files were converted for each specific program format using PGDSpider v.2.1.1.5 (Lischer & Excoffier, 2012).

To characterize each population of *L. sanguineus*, molecular diversity indexes (nucleotide diversity [θ_π] and nucleotide differences [θ_S]) were calculated in Arlequin (Excoffier & Lischer, 2010) using 1,000 simulations. Arlequin was also used to calculate pairwise $F_{ST,}$ global and population specific F_{IS} (Weir & Cockerham, 1984), and AMOVA. The multi-loci estimates of observed (H_o) and expected (H_e) heterozygosity, and global F_{ST} (Weir & Cockerham, 1984) were calculated in the package Adegenet (Jombart, 2008) in R software (R core time 2013). Due to the uneven number of specimens from *N. berthalutzae* and *N. pamelaroeae* collected in the different sample sites, some analysis of intra and interpopulation indexes could not be performed for these species. For those, instead of sample sites, the Structure clusters (see below) were used as populations. Only molecular diversity indexes (nucleotide diversity [θ_π] and nucleotide differences [θ_S]), global and pairwise F_{ST} (Weir & Cockerham, 1984) were calculated.

The Structure software (Pritchard et al., 2000) was used to infer the genetic clusters that compose each species populations. For this, analysis with the complete SNP dataset, and outlier SNPs datasets (see below) were run taking values of K from 1 to 2X the number of sampling sites to account for possible microgeographic structure with default settings, except for lambda and alpha values. Alpha value was set to one in all runs and lambda value was first estimated in one run with all other

values fixed. The estimated lambda value was fixed for 10 other runs. The results from the 10 iterations with fixed lambda values were compacted and analyzed in the R package Structure Harvester (Earl & vonHoldt, 2012). This package compares the likelihood of each K, and selects the most suitable K. The selected K was used to generate a bar plot showing the most likely genetic ancestry of each individual in the R package PopHelper (Francis, 2017).

The whole dataset was also used in the Adegenet package as input for Mantel test (Mantel, 1967), with 1,000 simulations, to identify possible isolation-by-distance (IBD) through the correlation between geographical and genetic distances for *L. sanguineus*. For this, each collection site was considerate as a different population. The genetic structure was inferred using the pairwise F_{ST} (Weir & Cockerham, 1984) between sites. The geographic distance was calculated using Google Maps (Google Coorporation), as the smallest distance between two sites in a straight line.

To identify SNPs putatively under selection, we used BayeScan (Foll & Gaggiotti, 2008) and an association between environmental variables and SNP frequency with the LFMM (latent factor mixed models tests) (Frichot et al., 2013; Rellstab et al., 2015) function of the R package LEA (Frichot & François, 2015), with the whole SNPs dataset, except for *N. pamelaroeae* populations which we used only LFMM due to the low number of specimens collected. In BayeScan, we ran 20 pilot runs of 50,000 iterations followed by 100,000 simulations with prior odd of 10 and 5% false discovery rate (FDR). Each location was considered as a different population. In LEA we used both the genetic clusters defined by Structure ±2 and the number of sampling locations ±2. Only the resulting SNPs in all analyses with different K values were considered to be associated to environmental variables. The environmental variables from each sampling site were downloaded from the

161

databases WordClim – 19 variables (available at: www.worldclim.org) and Bio-Oracle – 25 variables (available at: https://bio-oracle.org/) (Table 5). Using the variables of each database separately, we performed a PCA (*Principal Component Analyses*), and the four variables, from each database, that better explained the species distribution were selected for the association analyses with the LFMM function. We only selected four variables from each database due to computational constrains.

The outlier SNPs identified by both softwares were pooled together and used as input for population structure analyzes in Structure. The same was done with the neutral dataset, the SNPs not identified by any of the programs. These datasets were used as input for PCA using the package Adegenet. The outlier SNPs were also blasted against the NT and NR databases from NCBI. The identified proteins and non-coding RNAs were searched at the UniProt database (https://www.uniprot.org) to identify their function. In addition, the outlier loci from *N. berthalutzae* and *L. sanguineus* were also blasted against an annotated transcriptome of each species (Sonoda *et al.* in prep.).

The recent migration matrix and demographic scenarios were estimated in Fastsimcoal v.2.6 (Excoffier & Foll, 2011) based on the whole SNPs matrix. The input used was the minor allele site frequency spectrum parameter (SFS), computed with Arlequin with 1,000 bootstraps. From the SFS file, the recent migration rates were estimated either using each sample site (for *L. sanguineus* populations), and the Structure clusters as a population for *Nemertopsis* due to its low sample size. For migration estimates, we use 10 independent replicates, each including 40 estimation loops with 60,000 coalescence simulations and assuming current migration between

all pair of populations. For demographic scenarios, we estimated the origin of population expansion for all three species in Fastsimcoal. In this case, the Structure clusters were compared to the geographic distribution and from these we assumed populations (see more in the results section). The source of population expansion was simulated as each cluster. For *L. sanguineus* we ran four different models, for *N. berthalutzae* and *N. pamelaroeae* we ran two different models. For each model, 50 independent replicates, each including 40 estimation loops with 300,000 coalescence simulations, were performed. The probability of each model given the observed data was determined based on both the maximum likelihood value and AIC.

The recent migration pattern was also evaluated through the Estimated Effective Migration Surface (EEMS) software. For this, we ran two chains of 9,000,000 interactions, with a burn-in of 200,000 and thinning of 20,000. The convergence graph and other graphic output files were generate with rEEMSplots (Petkova et al., 2016) in R.

RESULTS

Prospection of SNPs and intrapopulation results

Sequenced libraries resulted in 159,706,948 raw reads for *Lineus sanguineus*, ranging between 316,950 – 7,010,816 reads per sample; 77,350,516 for *Nemertopsis berthalutzae*, ranging between 625,322 – 8,591,471 reads per sample; 46,271,716 for *Nemertopsis pamelaroeae*, with reads per sample ranging between 542,132 – 8,988,290 (Table S1). The final read counting, after several filtering and clustering steps, are: 773,611 reads for *L. sanguineus*, 216,564 for *N. berthalutzae*, and 130,442 for *N. pamelaroeae*. From those, we prospected 1,758 SNPs from 53 specimens of *L.

sanguineus, 2,601 SNPs from 19 specimens of *N. berthalutzae,* and 1,204 SNPs from 16 specimens of *N. pamelaroeae*. The prospected SNPs revealed that three out of seven populations had observed heterozygosity values significant higher than expected among *L. sanguineus* samples (Table 2). The mean H_O for *L. sanguineus*, *N. berthalutzae* and *N. pamelaroeae* is 0.174, 0.185 and 0.108, respectively.

Table 2. Number of samples (n), observed heterozygosity (H_O), expected heterozygosity (H_E), and fixation coefficient (F_{IS}) values across all loci per sample for all three species. Geographic region between parentheses. *p value < 0.001

Location	n	H_O	H_E	*t* test	F_{IS}
Lineus sanguineus					
Sabiaguaba (NE)	7	0.178	0.130	17.656*	-0.944
Boca da Barra (SE)	8	0.119	0.135	-0.922	-0.124
Itaipu (SE)	7	0.227	0.204	5.049*	-0.117
Boa Viagem (SE)	9	0.183	0.219	-7.85*	0.101
Araçá (SE)	12	0.139	0.174	-7.801*	0.094
Praia do Costão (SE)	10	0.201	0.157	10.977*	-0.465
Nemertopsis berthalutzae					
Sabiaguaba (NE)	3	0.097	0.102	-1.759	0.186
Araçá (SE)	13	0.120	0.165	-23.617*	0.279*
Nemertopsis pamelaroeae					
Jabaquara (SE)	8	0.177	0.345	-12.18*	0.451
Igararecê (SE)	4	0.192	0.287	-29.26*	0.077

Population structure with all loci

The Structure results from the matrices with all 1758 SNPs apportioned specimens of *L. sanguineus* in nine genetic clusters, despite they were sampled in only six different sites ($\Delta K = 255.17$; Fig. 2A). Most clusters correspond to each geographic location, except for four specimens that have similar contributions from all genetic clusters. In addition, one cluster is most common in two sites, which are just about 10 km apart (Boa Viagem and Itaipu, SE). The PCA results grouped most specimens from all Southeast locations in one group, but three specimens of Boa Viagem and five from Itaipu are in a second group, while six from Araçá are in a

third group. The specimens from the Northeast (Sabiaguaba) are apart from all other specimens in a fourth group (Fig. 2B). The pairwise F_{ST} show high and significant values between all pairs (0.242 – 0.523, $p < 0.05$), except for the Itaipu – Boa Viagem pair (0.001, $p = 0.396$) (Table 3). The AMOVA results show most of the variation is within individuals, followed by the variation within groups (Table 4). The global value of F_{ST} is high and significant, while the global value of F_{IS} is negative and not significant (Table 4).

Table 3. Pairwise F_{ST} (below diagonal) and p-values (above diagonal) between all sampling sites with all loci considered for *Lineus sanguineus* specimens. p value < 0.001*

	Sabiaguaba	Boca da Barra	Itaipu	Boa Viagem	Araca
Boca da Barra (SE)	0.523*				
Itaipu (SE)	0.429*	0.409*			
Boa Viagem (SE)	0.370*	0.363*	0.001		
Araca (SE)	0.469*	0.310*	0.321*	0.287*	
Praia do Costão (SE)	0.429*	0.327*	0.284*	0.243*	0.267*

Table 4. AMOVA results and global fixation indexes with all loci considered for *Lineus sanguineus.*

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	Global fixation indexes
Among groups	1	191.494	-1.67196	-4.22	$F_{ST} = $ **0.358**
Among populations within groups	4	1125.451	14.73564	37.22	$F_{IS} = -0.157$
Within populations	47	1051.187	-4.15678	-10.50	$F_{SC} = $ **0.357**
Within individuals	53	1626.000	30.67925	77.50	-

Table 5. Environmental variables selected for LFMM association. The number of SNPs associated and loci identified with each variable is presented.
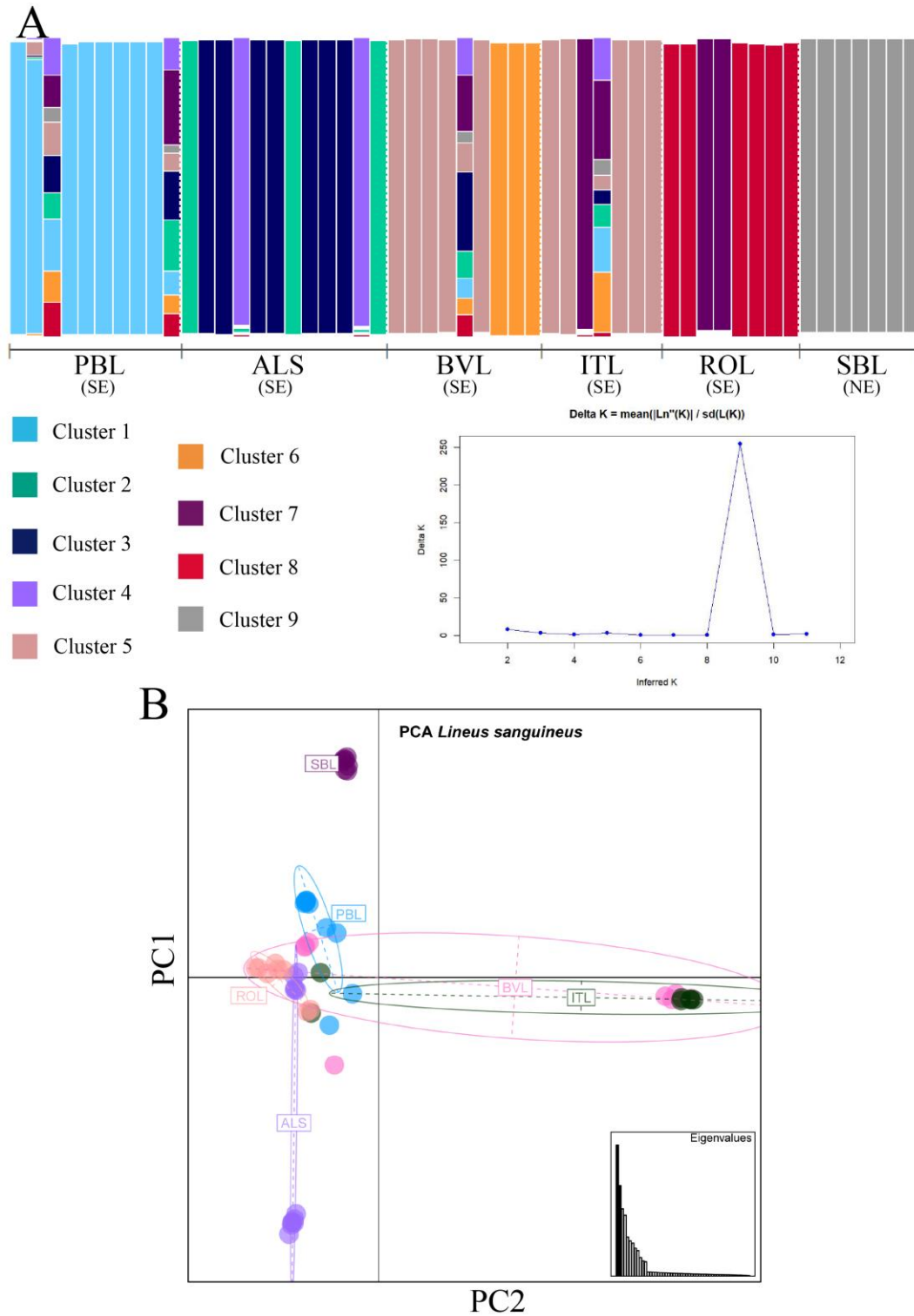
Lineus sanguineus

Source database	Variable #	Variable Name	# SNPs associated	# loci identified
Bio-Oracle	Variable 1	Dissolved oxygen concentration, maximum at min depth	161 SNPs	13 loci
	Variable 2	Chlorophyll A, maximum at sea surface	32 SNPs	4 loci
	Variable 3	Chlorophyll A, mean at sea surface	35 SNPs	4 loci
WorldClim	Variable 4	Minimum Temperature of Coldest Month	149 SNPs	13 loci
	Variable 5	Mean Temperature of Coldest Quarter	152 SNPs	14 loci
	Variable 6	Precipitation of Wettest Month	63 SNPs	6 loci
	Variable 7	Precipitation of Wettest Quarter	32 SNPs	2 loci
Nemertopsis berthalutzae				
Source database	Variable #	Variable Name	# SNPs associated	# loci identified
Bio-Oracle	Variable 1	Sea water temperature, range at min depth	4 SNPs	1 loci
	Variable 2	Chlorophyll A, mean at sea surface	268 SNPs	7 loci
	Variable 3	Dissolved oxygen concentration, longterm max at min depth	3 SNPs	0
	Variable 4	Salinity, minimum at sea surface	473 SNPs	23 loci
WorldClim	Variable 5	Mean Temperature of Coldest Quarter	8 SNPs	0
	Variable 6	Precipitation of Coldest Quarter	40 SNPs	0
	Variable 7	Annual Mean Temperature	8 SNPs	0
	Variable 8	Max Temperature of Warmest Month	19 SNPs	0

Structure separated the specimens of *N. berthalutzae* in three clusters, despite

they were collected in four different locations (Fig. 3), one predominant between

specimens from Northeast, the second most common between specimens of the Southeast and the third present only in a few specimens from the Southeast (Fig 3A). However, the PCA grouped all specimens in only one group with a few outlier individuals from Araçá and Mambucaba, both in the Southeast (Fig. 3B). Taking these results and the geographic proximity between the sampling sites, we delimitated two populations, the first encompassing specimens from Sabiaguaba (NE) and Itaoca (SE), having a $\theta_S = 99.89$ (s.d. = 43.17) and $\theta_\pi = 91.96$ (s.d. = 50.56), while the second with specimens form Praia das Goiabas (SE) and Araçá (SE), having $\theta_S = 47.46$ (s.d. = 14.85) and $\theta_\pi = 32.09$ (s.d. = 16.03). The pairwise F_{ST} value was low (0.0574), but significant ($p = 0.036$).

The Structure results of *N. pamelaroeae* indicate the presence of four genetic clusters. All clusters seem to contribute to the genetic background of all specimens, but two from Igararecê (SE). These two specimens belong exclusively to one cluster that has just a minimum contribution in the genetic background of the other specimens (Fig. 4A). PCA's results show a similar pattern with the two specimens from Igararecê very distant from the others, while the other specimens are separated mostly by PC1 (Fig. 4B). With these results, we considerate two clusters: the first with specimens only from Igararecê (SE), having $\theta_S = 97.58$ (s.d. = 42.17) and $\theta_\pi = 117.68$ (s.d. = 64.59), and the second with all other specimens, having $\theta_S = 1.65$ (s.d. = 0.83) and $\theta_\pi = 1.70$ (s.d. = 1.15). The pairwise F_{ST} was high (0.293) and significant ($p = 0.018$).
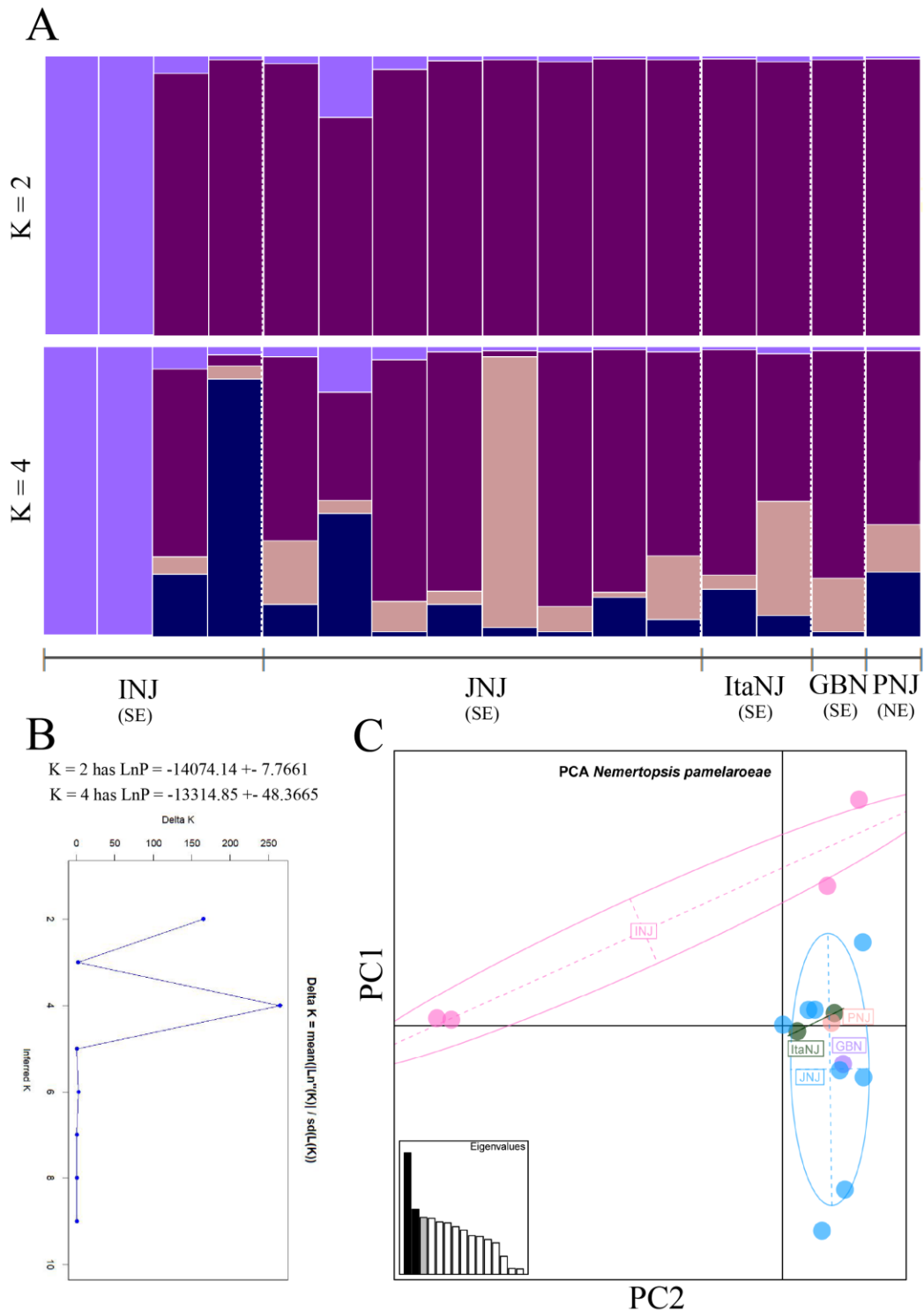
Figure 2. Clustering results for 1,758 SNPs from *L. sanguineus* individuals. A: Structure results showing the best K as indicated by ΔK values (presented in lower half); B: PCA results. Each circle corresponds to one individual, different colors represent different sample sites.
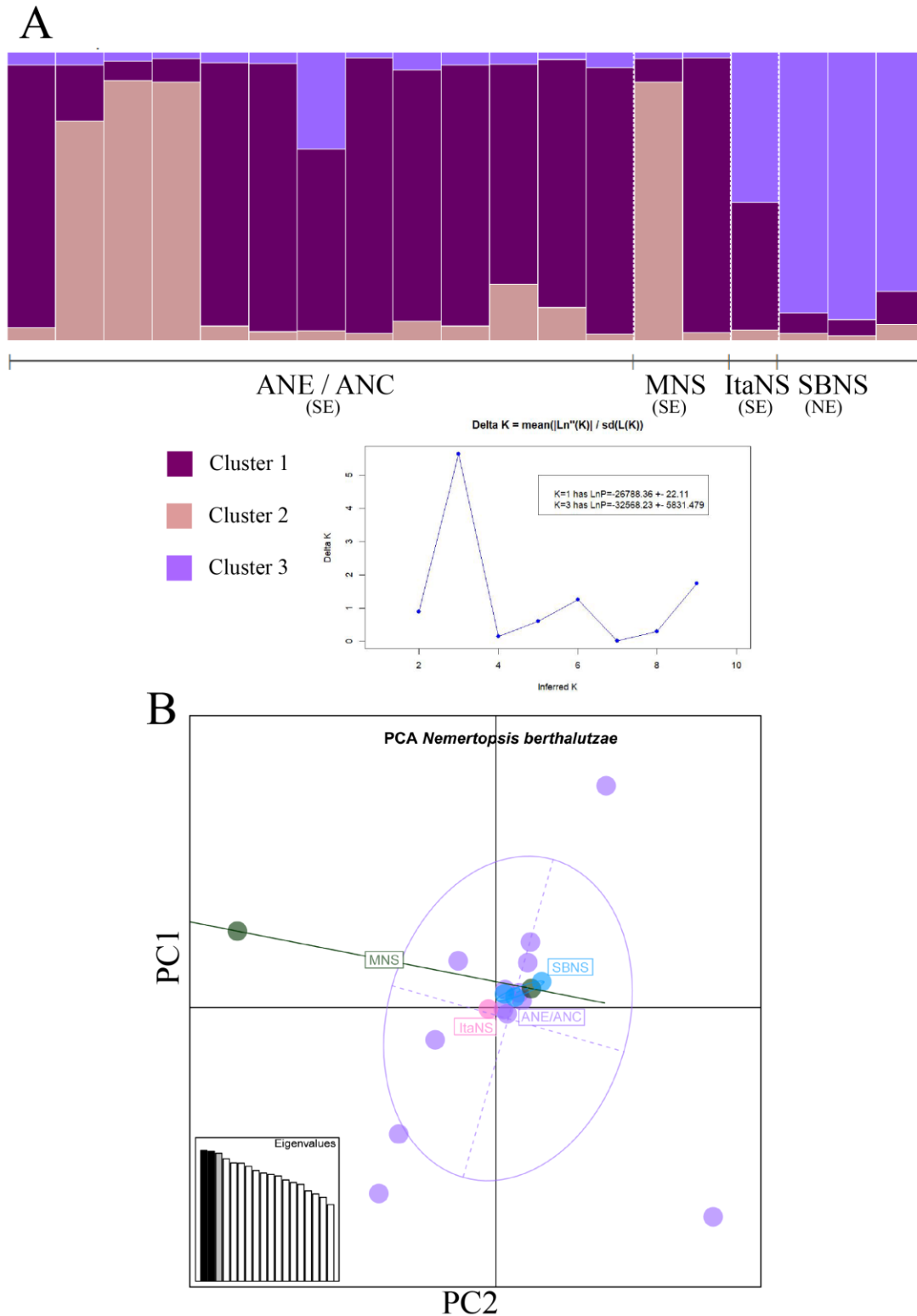
Figure 3. Clustering results for 2,601 SNPs from *N. berthalutzae* individuals. A: Structure results showing the best K as indicated by ΔK values (presented in lower half, note insert with likelihoods for K=1 and K=3); B: PCA results. Each circle corresponds to one individual, different colors represent different sample sites.

Figure 4. Clustering results for 1,204 SNPs from *N. pamelaroeae* individuals. A: Structure results showing K=2 and K=4, as explained in main text; B: ΔK values from Structure Harvester results, note insert with likelihoods for K=2 and K=4; C: PCA results. Each circle corresponds to one individual, different colors represent different sample sites.

Migration pattern estimates and demographic history

According to the EEMS results none of the three species present any regions of higher than average migration along the Brazilian coast, only regions of low than average migration located in the southeast region (Fig. 5 A, C, E). The EEMS estimates of genetic diversity (Fig. 5 B, D, F) also show similar patterns for the three species, with two regions of low than average diversity, one in the northeast region and another in the southeast region. However, the site in the southeast region of low diversity for *L. sanguineus* is located in the southernmost area, while for the two *Nemertopsis* species, this area is located near the northernmost are of the southeast. The migration estimates between the populations of *L. sanguineus* generated in the Fastsimcoal software show heterogeneous rates of migration between the populations, and no specific direction or distinguishable pattern was observed (Fig. 6A). For *N. berthalutzae,* the migration estimates between the Northeast (SBNS + ItaNS) and Southeast (MNS + ANC + ANE) show 10 times more migrants from Southeast to Northeast than the other way around (Fig. 6B). The migration rates are; however, lower than what found for *L. sanguineus* populations. The estimates for *N. pamelaroeae* populations show even lower migration rates, with a higher migration coming from Igararecê (Fig. 6C). The Mantel test shows a positive and significant correlation between geographic and genetic distance for the populations of *L. sanguineus* (R = 0.636; *p* = 0.031). While is non-significant for the two *Nemertopsis* species.

The Fastsimcoal results are not very clear about the demographic past of all three species. Due to software constrains, we used the PCA results and geographical information to delimitate four genetic populations for *L. sanguineus* populations. The first encompassing the specimens from Sabiaguaba (NE), the second with specimens

from Itaipu and Boa Viagem (SE), the third comprises specimens from Boca da Barra and Araçá (SE), and the fourth with specimens from Praia do Costão (SE). We tested these four clusters as the possible source population from where the Brazilian populations expanded (Fig. 7). The AIC and likelihood values show different scenarios. Likelihood values cannot distinguish well between the Southeast expansion scenarios, while AIC points the expansion from NE as the most likely scenario (Fig. 7). The estimates for *N. berthalutzae* suggest an old expansion from Northeast (Sabiaguaba + Itaoca), with a divergence time older than 300 thousands generations (Fig. 8). For *N. pamelaroeae* populations Fastsimcoal also cannot distinguish between scenarios, with likelihood without separate between the scenarios, and AIC values pointing to an expansion from Igararecê (SE), but with really low difference to the other scenario (Fig. 9). The time of divergence in both cases is also very old, being more than 150 thousands generations for Igararecê expansion.
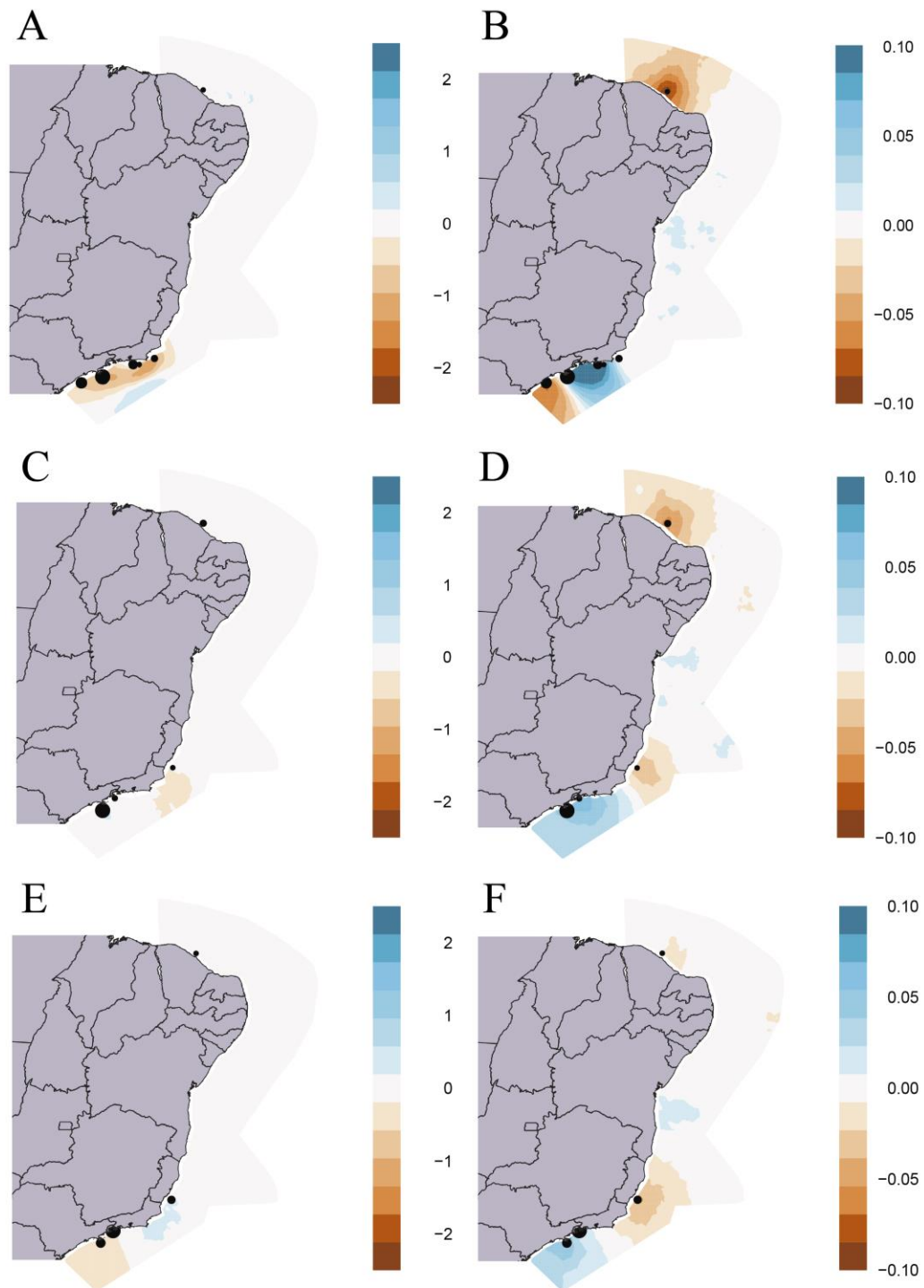
Figure 5. Migration and diversity surface estimates from EEMS. A: migration estimate for *L. sanguineus* populations; B: diversity estimate for *L. sanguineus* populations; C: migration estimate for *N. berthalutzae* populations; D: diversity estimate for *N. berthalutzae* populations; E: migration estimate for *N. pamelaroeae* populations; F: diversity estimate for *N. pamelaroeae* populations.
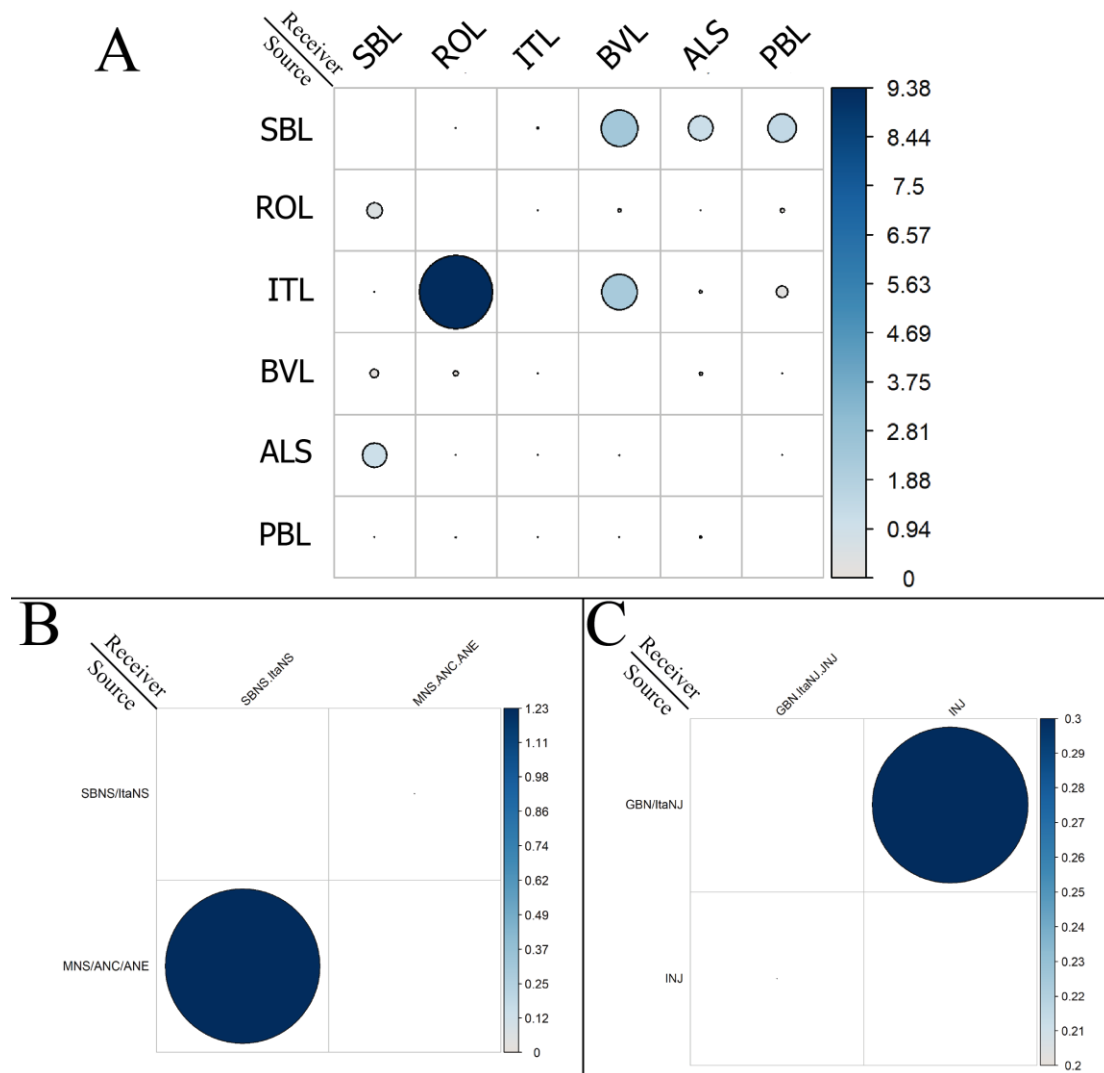
Figure 6. Migration rates per generation as estimated by Fastsimcoal. A: estimate for *L. sanguineus* populations; B: estimate for *N. berthalutzae* populations; C: estimate for *N. pamelaroeae* populations. Circle sizes are proportional to migrant numbers.
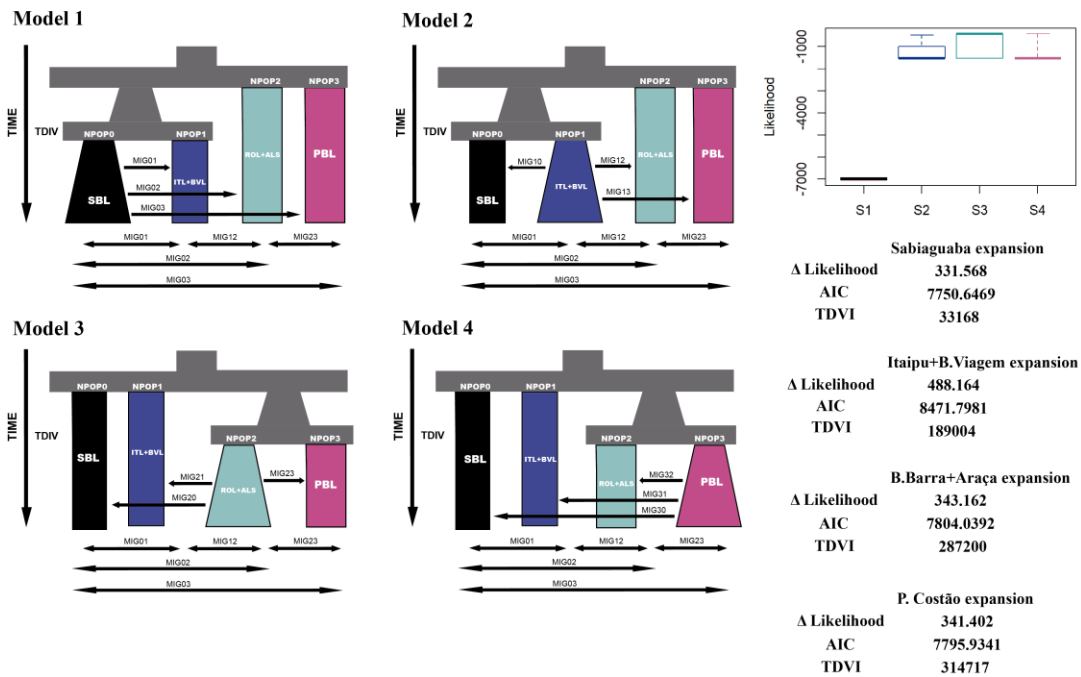
Figure 7. *Lineus sanguineus* demographic models simulated in Fastsimcoal (left side) and their results (right side). Models are presented backward in time, as represented by the black arrow on the left. The table on the right side shows the comparison between the results from each model. Boxplot presents the individual likelihoods of each model.
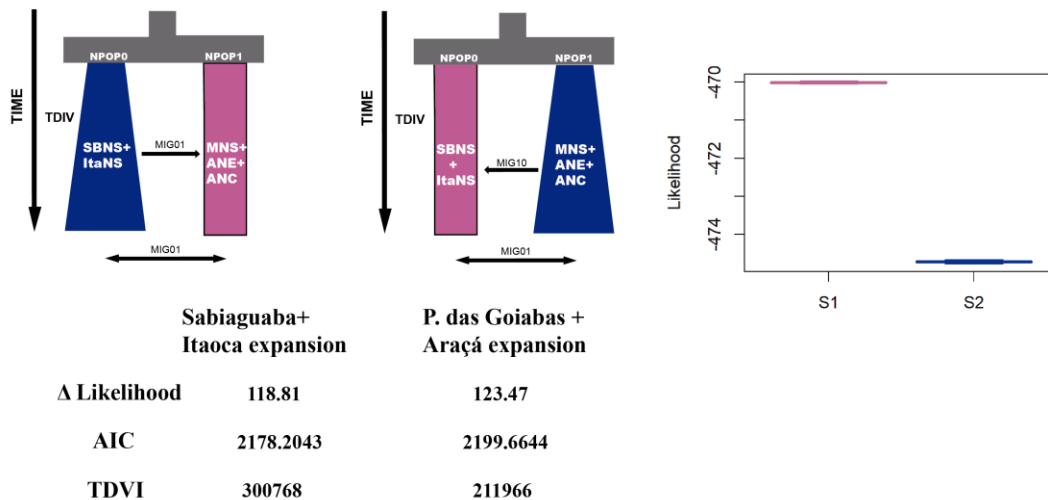
Figure 8. *Nemertopsis berthalutzae* demographic models simulated in Fastsimcoal (upper half) and their results (lower half). Models are presented backward in time, as represented by the black arrow on the left. The table bellows shows the comparison between the results from each model. Boxplot presents the individual likelihoods of each model.

175
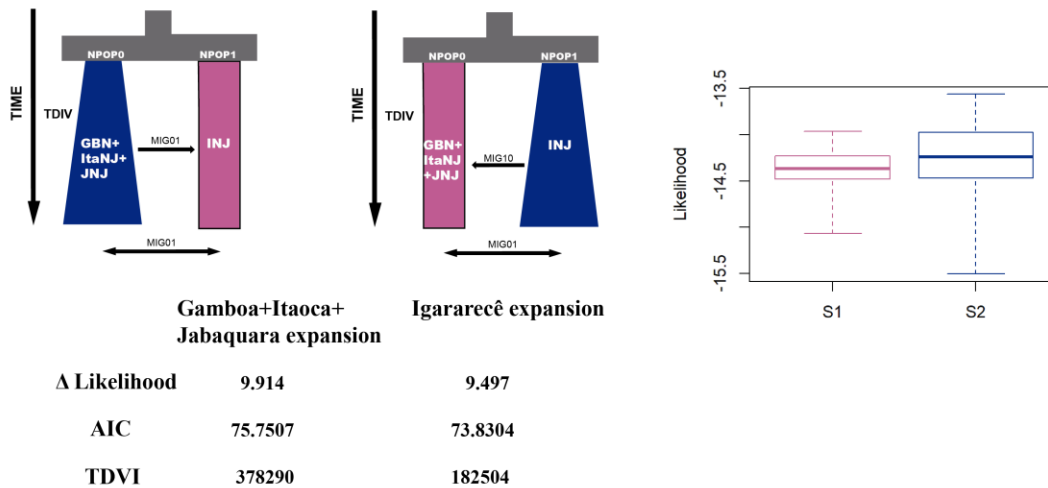
Figure 9. *Nemertopsis pamelaroeae* demographic models simulated in Fastsimcoal (upper half) and their results (lower half). Models are presented backward in time, as represented by the black arrow on the left. The table bellows shows the comparison between the results from each model. Boxplot presents the individual likelihoods of each model.

SNPs putatively under selection

For the populations of *L. sanguineus*, BayeScan identified 27 outlier SNPs (Fig. S1), while LEA identified 247 SNPs in 179 loci correlated to seven environmental variables (Table 5; Fig. S2). This species only had seven environmental variables selected in the PCA because the variables from Bio-Oracle that better explained their distribution were all related to chlorophyll A (Table S2), and select different measurements of the same variable did not seem to produce any additional information. LEA was not able to identify any outlier SNPs among the *N. pamelaroeae* populations. *Nemertopsis berthalutzae*, in its turn, also had no outlier SNPs identified by BayeScan, while LEA presented 519 SNPs in 171 loci, related to eight environmental variables (Table 5; Fig. S3).

The Blast search for *L. sanguineus* loci had no results against both the nt and nr databases, only against the transcriptome. From the 274 loci searched, 78 were identified (Table S3). They are related to all seven environmental variables and are

hits to genes involved with activation of transcription processes (*KDM5A, TAF4, YETS2, DPOLA, NFYA, SNPC4, RNF10, DHX34, HDGR2*); neural response to external stimulus (*NEC2, CTRO, GRM2*); response to DNA damage and stress (*TMP2L, NRDE2, USP9X, SACS, CRHBP*); and immune system (*CD9, BLNK*). From the 27 loci indicated by BayeScan as under positive selection, only one had hit during the Blast searches, to Ankyrin-2 gene, which is related to microtubule cytoskeleton organization.

The Blast searches against the nr database also had no results for *N. berthalutzae* loci, while nine loci (related to the environmental variables 1 and 4) had hits against the nt database, identifying six genes *SMAD5, SMAD9, PSMA3, TMEM184B, PDE8B* and *yy1*, which were also identified in the searches against our in house transcriptome. This search showed hits to loci related to environmental variables 1, 2 and 4 (Table 5), identifying 22 loci. The genes are involved with activation of transcription processes (*CSRN3, UBIP1, SMAD5, EIF3G*); adipose tissues biogenesis (*SPTC2*); response to DNA damage (T184B); and immune response (*TENS1, MACF1, DYH7*), besides genes involved with the normal cell cycle.

The outlier SNPs were also used as input to investigate population structure, to evaluate any possible structuring related to the environmental variables. The results from *L. sanguineus* individuals from PCA and Structure show a similar pattern with three clusters (Fig. S4). Some specimens from Araçá, Boa Viagem and Itaipu, however, are discordant between the two analyses. The results PCA from *N. berthalutzae* show the same results as the analysis with all SNPs. The results from Structure harvester, however, show the most probable K value as seven, but the

Structure results with K=7 show a very strange distribution, without any distinguishable pattern (Fig. S5).

DISCUSSION

This is the first comparative study using seascape genomics to understand the population dynamics between intertidal species with different reproductive modes from the Brazilian coast. We found that, as expected, the probably partial asexual species shows more structured populations than the species with planktotrophic development. Nonetheless, *Lineus sanguineus* presented higher migration rates compared to the hoplonemertean species (Fig. 6). Evidences of environmental adaptation seem to be related to similar conditions, especially the temperature and precipitation, regardless of the different developmental types.

Comparative population genomics

Only *Nemertopsis berthalutzae* Araçá population presented significant F_{IS}. Populations of *N. berthalutzae* and *N. pamelaroeae* show a tendency towards positive values of F_{IS}, while *L. sanguineus* presents a tendency towards negative values. The presence of heterozygosity deficiency, denoted by positive values of F_{IS}, can be related to many factors, being the most common for invertebrate species, mixing of subpopulations composed by different cohorts, the Wahlund effect (Castric et al., 2002; Ganapathi et al., 2012; Haeringen et al., 1999), which is the most probable for two *Nemertopsis* species. The heterozygosity excess can be, in most cases, related to balancing selection, recent bottleneck, hybrid origin of asexual lineages or sequence divergence during long-term asexuality, the Meselson effect (Ament-Velásquez et al., 2016; Ceplitis, 2003; Delmotte et al., 2002). Being the latter the most probable cause for the F_{IS} values found for *L. sanguineus* populations.

As expected, the Structure analyses also show a discordant pattern between *L. sanguineus* and the two *Nemertopsis* species. The results for *L. sanguineus* show nine clusters, each roughly correspondent to a collection site, except for the two sites less than five kilometers apart (Itaipu and Boa Viagem) that share the same cluster, while the PCA indicates only four clusters, not compatible with the collection sites. This discrepancy is likely a result of the missing data present in our dataset, because PCA results are commonly affected by amounts of missing data higher than 10% (Yi & Latch, 2021). When the missing data shows a non-random pattern of missing data the PCA plots can have a bias towards the origin of the PCA, while the random missing data reduce data variation and analytical power, but do not seem to cause biased results, except for more loose genetic clusters. Therefore, the population structure seen in PCA plots need to be complemented by other analyses if the dataset presents more than 10% of missing data (Yi & Latch, 2021), like the genotype assignment approach in the Structure software. The results of PCA and Structure also show some discrepancy in the two *Nemertopsis* species. Nonetheless, in the case of *N. berthalutzae,* despite the lower than ideal variation explained by PCs 1 and 2, this discordance is most likely an effect of the Evanno method to choose the best K. This method calculates the most likely value of K using the Ln value of K-1, consequently K=1 cannot be included in this calculus. However, one can notice when the best K=1 using the variance values (Earl & vonHoldt, 2012; Evanno et al., 2005). This because, once the real value of K is reached the variance between runs increases, as seen in *N. berthalutzae* Structure results (Fig. 3). This increased variance is also present in the Structure results of *N. pamelaroeae*, where the ΔK value indicates K=4 as the best K, however the K=2 has a lower variance indicating that this might be the true value of K. This is also the number of clusters found in the PCA analysis, with the same

groups formed. Nonetheless, one of the groups has only two specimens, despite this two share the same sampling site with other two specimens that are in a different cluster. *Nemertopsis pamelaroeae* individuals usually live in fouling communities, and because of it they are more prone to be carried in ship hulls (Carlton, 1985). Since the sampling site where these animals were collected is marina, it is very likely that these two genetically different individuals are from an alien population.

Migration pattern estimates

The migration estimates from EEMS show average migration in most of the coast for all three species. However, for *L. sanguineus* the Southeast populations seem to have a low than average migration between each other, what is further confirmed by the Fastsimcoal estimates, where most of the migration comes from the Northeast (SBL). Since such estimates are result of shared diversity between populations (Excoffier & Foll, 2011; Petkova et al., 2016), this can either be a result of encysted pieces or larvae floating on the surface currents, or carried on the fouling community on ship hulls. Nonetheless, our population results indicate the presence of sexual reproduction, as seen in Ament-Velasquez et al. (2016), what points towards the connection of populations through larval dispersion, which is also the most common way for partially asexual species to disperse for long distances (Jackson, 1986). In addition, the presence of lower than average migration between the sites in the Southeast might be an indication of the presence of some asexual reproduction in these sites, what is also indicated by the PCA results where specimens from the same sampling site belong to different clusters (Brandt et al., 2021). *Lineus sanguineus* is frequently cited in the literature as a fissiparous species (Gontcharoff, 1950; Moretto & Brancato, 1997; Runnels, 2013). In addition, previous comparative transcriptome assessment of *L. sanguineus* and its two closest congeners *Lineus lacteus* and *Lineus*

pseudolacteus found that *L. pseudolacteus* is a probable asexual hybrid between *L. lacteus* and *L. sanguineus,* being *L. lacteus* a sexual species and *L. sanguineus* a species with sexual and asexual reproduction (see Ament-Velásquez et al., 2016). Therefore, it is likely that the excess of heterozygosity found here is a product of the Meselson effect, but since we did not detect any significant F_{IS} values, we can only assume that the sexual reproduction might be playing a role in these populations.

The migration estimates for *N. berthalutzae* and *N. pamelaroeae* from Fastsimcoal show very low rates with one major source of migrants. However, the migrant flow between *N. berthalutzae* populations is 10x stronger from the Northeast than the other way around. Such pattern is commonly seeing for coastal species that disperse through the southward flowing Brazil Current, such as the bivalve *Anomalocardia brasiliana*, groupers from the genus *Mycteroperca* and the polychaete *Perinereis ponteni* (Arruda et al., 2009; D'Agostini et al., 2015; Mendes *et al.* subm.). Three main surface currents flow along the Brazilian coast, two branches of the South Equatorial Current that reaches the coast at the Northeast region, the North Brazil Current (flowing northwards) and the Brazil Current (flowing southwards), and the Malvinas Current, a branch of the Circumpolar Current (flowing northwards) (N. Martins et al., 2021). These currents also delimitate the marine Ecoregions present in the coast, the North Brazil Shelf, the Tropical Southwestern Atlantic, and the Worm Temperate Southwestern Atlantic (Spalding et al., 2007).

The diversity patterns estimated in EEMS also show a similar pattern for all three species, with the site in the Northeast (SBL, SBNS and PNJ) showing less than average diversity, while a site in the Southeast has higher than average diversity. Once again this seems to be a result of the sampling bias, where sites with more sampled animals showing higher diversity rates. These diversity rates, nevertheless,

contrast with the Fastsimcoal results that indicate the Northeast populations as source for the expansion in *L. sanguineus* and *N. berthalutzae*. The Fastsimcoal was not able to clearly distinct the source populations in both cases. Therefore, it is possible that demographic results were affected by the uneven sampling in the different locations.

Evidences of adaptive selection

No signs of adaptive selection were detect among the populations of *N. pamelaroeae*. However, this cannot be interpreted as absence of selection among these populations. BayeScan software has very conservative algorithm with a lower detection power when selection effects are not intense or when only a small number of populations is analyzed (De Mita et al., 2013; Tigano et al., 2017). In addition, the uneven number of specimens per sampling site and the chosen value of K can affect the association test performed in the LFMM package (Caye et al., 2019; Frichot et al., 2013; Frichot & François, 2015; McVean, 2009).

The presence of local adaptation with genes related to immune and DNA damage response in both *L. sanguineus* and *N. berthalutzae* might be an indication of anthropogenic contaminants in the sampled areas. Most sampling sites were located near estuaries in populated areas, and such areas along the Brazilian coast are known for contain sewage contamination, as well as contaminants from fossil fuels (Araújo et al., 2020; Bertrand et al., 2016; Frena et al., 2016; Martins et al., 2012). Furthermore, the contaminants can decrease the oxygen concentration in the water and facilitate the proliferation of pathogenic microorganisms (Meyer-Reil & Köster, 2000), as well as carry mutagenic substances causing DNA damage and infections. This can also be the cause for so many genes involved with activation of transcription processes. The genes identified here are not yet characterized in invertebrate species. Nonetheless, most of them seem to be related to heterochromatin activation and

inactivation, what is related to maintenance of genomic stability (Allshire & Madhani, 2018) and can, therefore, also be involved in response to environmental stress.

The environmental variables related to the local adaptation are also similar for the two species, being directly or indirectly related to temperature and precipitation. This can be an effect of the large latitudinal gradient covered in our sampling. This area has great differences in temperature and precipitation during the year (Reboita et al., 2015), affecting marine and terrestrial populations. Interestingly the different rates of gene flow of the two species were not an important factor in shaping the local adaptation, demonstrating how strong these environmental variables are affecting these populations. This brings attention to the possible effect of environmental change will have in these populations, as well as all other that live in similar environments.

staff of CEBIMar-USP for providing the essential laboratory facilities and logistics for this study.

REFERENCES

Allshire, R. C., & Madhani, H. D. (2018). Ten principles of heterochromatin formation and function. *Nature Reviews Molecular Cell Biology*, *19*(4), 229–244. https://doi.org/10.1038/nrm.2017.119

Ament-Velásquez, S. L., Figuet, E., Ballenghien, M., Zattara, E. E., Norenburg, J. L., Fernández-Álvarez, F. A., Bierne, J., Bierne, N., & Galtier, N. (2016). Population genomics of sexual and asexual lineages in fissiparous ribbon worms (Lineus, Nemertea): Hybridization, polyploidy and the Meselson effect. *Molecular Ecology*, *25*(14), 3356–3369. https://doi.org/10.1111/mec.13717

Andrade, S. C. S., Norenburg, J. L., & Solferini, V. N. (2011). Worms without borders: Genetic diversity patterns in four Brazilian *Ototyphlonemertes* species (Nemertea, Hoplonemertea). *Marine Biology*, *158*(9), 2109–2124. https://doi.org/10.1007/s00227-011-1718-3

Andrade, S. C. S., Strand, M., Schwartz, M., Chen, H., Kajihara, H., von Döhren, J., Sun, S., Junoy, J., Thiel, M., Norenburg, J. L., Turbeville, J. M., Giribet, G., & Sundberg, P. (2012). Disentangling ribbon worm relationships: Multi-locus analysis supports traditional classification of the phylum Nemertea. *Cladistics*, *28*(2), 141–159. https://doi.org/10.1111/j.1096-0031.2011.00376.x

Araújo, M. P., Hamacher, C., de Oliveira Farias, C., Martinho, P., de Oliveira Chaves, F., & Gomes Soares, M. L. (2020). Assessment of Brazilian mangroves hydrocarbon contamination from a latitudinal perspective. *Marine Pollution Bulletin*, *150*, 110673. https://doi.org/10.1016/j.marpolbul.2019.110673

Becker, B. J., Levin, L. A., Fodrie, F. J., & McMillan, P. A. (2007). Complex larval connectivity patterns among marine invertebrate populations. *Proceedings of the National Academy of Sciences*, *104*(9), 3267–3272. https://doi.org/10.1073/pnas.0611651104

Bertrand, G., Hirata, R., Pauwels, H., Cary, L., Petelet-Giraud, E., Chatton, E., Aquilina, L., Labasque, T., Martins, V., Montenegro, S., Batista, J., Aurouet, A., Santos, J., Bertolo, R., Picot, G., Franzen, M., Hochreutener, R., & Braibant, G. (2016). Groundwater contamination in coastal urban areas: Anthropogenic pressure and natural attenuation processes. Example of Recife (PE State, NE Brazil). *Journal of Contaminant Hydrology*, *192*, 165–180. https://doi.org/10.1016/j.jconhyd.2016.07.008

Bird, A. M., von Dassow, G., & Maslakova, S. A. (2014). How the pilidium larva grows. *EvoDevo*, *5*(1), 13. https://doi.org/10.1186/2041-9139-5-13

Brandt, A., Tran Van, P., Bluhm, C., Anselmetti, Y., Dumas, Z., Figuet, E., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., Labédan, M., Maraun, M., Parker, D. J., Robinson-Rechavi, M., Schaefer, I., Simion, P., Scheu, S., Schwander, T., & Bast, J. (2021). Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proceedings of the National Academy of Sciences*, *118*(38), e2101485118. https://doi.org/10.1073/pnas.2101485118

Caplins, S. A., & Turbeville, J. M. (2011). The Occurrence of *Ramphogordius sanguineus* (Nemertea, Heteronemertea) in the Intertidal Zone of the Atlantic Coast of Virginia and New Observations on its Feeding Behavior. *Virginia Natural History Society 38*, 66–70.

Carlton, J. T. (1985). Transoceanic and interoceanic dispersal of coastal marine organisms: The biology of ballast water. *Transoceanic and interoceanic dispersal of coastal marine organisms: the biology of ballast water*, *23*, 313–371.

Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus fontinalis* Mitchill (Pisces, Salmonidae): A test of alternative hypotheses. *Heredity*, *89*(1), 27–35. https://doi.org/10.1038/sj.hdy.6800089

Caye, K., Jumentier, B., Lepeule, J., & François, O. (2019). LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies. *Molecular Biology and Evolution*, *36*(4), 852–860. https://doi.org/10.1093/molbev/msz008

Ceplitis, A. (2003). Coalescence times and the Meselson effect in asexual eukaryotes. *Genetical Research*, *82*(3), 183–190. https://doi.org/10.1017/S0016672303006487

Cowen, R. K., & Sponaugle, S. (2009). Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, *1*(1), 443–466. https://doi.org/10.1146/annurev.marine.010908.163757

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., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0062137

de Medeiros, B. A. S., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ*, *6*, e5089. https://doi.org/10.7717/peerj.5089

De Mita, Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*(5), 1383–1399. https://doi.org/10.1111/mec.12182

Delmotte, F., Leterme, N., Gauthier, J.-P., Rispe, C., & Simon, J.-C. (2002). Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Molecular Ecology*, *11*(4), 711–723. https://doi.org/10.1046/j.1365-294X.2002.01478.x

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bulletim*, *19*, 11–15.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A. R., & Overcast, I. (2020). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btz966

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

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. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., & Foll, M. (2011). fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, *27*(9), 1332–1334. https://doi.org/10.1093/bioinformatics/btr124

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

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. https://doi.org/10.1534/genetics.108.092221

Francis, R. M. (2017). pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32. https://doi.org/10.1111/1755-0998.12509

Frena, M., Bataglion, G. A., Tonietto, A. E., Eberlin, M. N., Alexandre, M. R., & Madureira, L. A. S. (2016). Assessment of anthropogenic contamination with sterol

markers in surface sediments of a tropical estuary (Itajaí-Açu, Brazil). *Science of The Total Environment*, *544*, 432–438. https://doi.org/10.1016/j.scitotenv.2015.11.137

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210X.12382

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution*, *30*(7), 1687–1699. https://doi.org/10.1093/molbev/mst063

Ganapathi, P., Rajendran, R., & Kathiravan, P. (2012). Detection of occurrence of a recent genetic bottleneck event in Indian hill cattle breed Bargur using microsatellite markers. *Tropical Animal Health and Production*, *44*(8), 2007–2013. https://doi.org/10.1007/s11250-012-0171-8

Gibson, R. (1995). Nemertean genera and species of the world: An annotated checklist of original names and description citations, synonyms, current taxonomic status, habitats and recorded zoogeographic distribution. *Journal of Natural History*, *29*(2), 271–561. https://doi.org/10.1080/00222939500770161

Gontcharoff, M. (1950). Sur La Reproduction Sexuee Chez Lineus-Sanguineus (lineus-ruber-beta). *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences*, *230*(2), 233–234.

Haeringen, W. V., Gwakisa, P. S., Mikko, S., Eythorsdottir, E., Holm, L. E., Olsaker, 1i, Outteridge, P., & Andersson, L. (1999). Heterozygosity excess at the cattle DRB locus revealed by large scale genotyping of two closely linked microsatellites. *Animal Genetics*, *30*(3), 169–176. https://doi.org/10.1046/j.1365-2052.1999.00436.x

Hellberg, M. E., Burton, R. S., Neigel, J. E., & Palumbi, S. R. (2002). Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science*, *70*(1), 18.

Ikenaga, J., Hookabe, N., Kohtsuka, H., Yoshida, M., & Kajihara, H. (2019). A Population Without Females: Males of *Baseodiscus delineatus* (Nemertea: Heteronemertea) Reproduce Asexually by Fragmentation. *Zoological Science*, *36*(4), 348–353. https://doi.org/10.2108/zs180203

Jackson, J. B. C. (1986). Modes of dispersal of clonal benthic invertebrates: Consequences for species' distributions and genetic structure of local populations. *BULLETIN OF MARINE SCIENCE*, *39*, 19.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kajihara, H., Chernyshev, A. V., Sun, S.-C., Sundberg, P., & Crandall, F. B. (2008). Checklist of Nemertean Genera and Species Published between 1995 and 2007. *Species Diversity*, *13*(4), 245–274. https://doi.org/10.12782/specdiv.13.245

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, *27*(2 Part 1), 209–220.

Martins, C. D. L., Arantes, N., Faveri, C., Batista, M. B., Oliveira, E. C., Pagliosa, P. R., Fonseca, A. L., Nunes, J. M. C., Chow, F., Pereira, S. B., & Horta, P. A. (2012). The impact of coastal urbanization on the structure of phytobenthic communities in southern Brazil. *Marine Pollution Bulletin*, *64*(4), 772–778. https://doi.org/10.1016/j.marpolbul.2012.01.031

Maslakova, S. A., & Hiebert, T. C. (2014). From trochophore to pilidium and back again—A larva's journey. *The International Journal of Developmental Biology*, *58*(6-7–8), 585–591. https://doi.org/10.1387/ijdb.140090sm

McVean, G. (2009). A Genealogical Interpretation of Principal Components Analysis. *PLOS Genetics*, *5*(10), e1000686. https://doi.org/10.1371/journal.pgen.1000686

Mendes, C. B., Norenburg, J. L., & Andrade, S. C. S. (2021). Species delimitation integrative approach reveals three new species in the *Nemertopsis bivittata* complex. *Invertebrate Systematics*. https://doi.org/10.1071/IS20048

Mendes, C. B., Norenburg, J. L., Solferini, V. N., & Andrade, S. C. S. (2018). Hidden diversity: Phylogeography of genus *Ototyphlonemertes* Diesing, 1863 (Ototyphlonemertidae: Hoplonemertea) reveals cryptic species and high diversity in Chilean populations. *PLOS ONE*, *13*(4), e0195833. https://doi.org/10.1371/journal.pone.0195833

Meyer-Reil, L.-A., & Köster, M. (2000). Eutrophication of Marine Waters: Effects on Benthic Microbial Communities. *Marine Pollution Bulletin*, *41*(1), 255–263. https://doi.org/10.1016/S0025-326X(00)00114-4

Moretto, H. J. A., & Brancato, C. L. (1997). The ovaries of a fissiparous heteronemertean, *Lineus bonaerensis*, from Argentina. *Hydrobiologia*, *365*(1/3), 129–134. https://doi.org/10.1023/A:1003138829001

Norenburg, J. L., Gibson, R., Herrera-Bachiller, A., & Strand, M. (2017). *World Nemertea Database*. World Register of Marine Species. http://www.marinespecies.org/nemertea

Nunes, J. de R. da S., Liu, S., Pértille, F., Perazza, C. A., Villela, P. M. S., de Almeida-Val, V. M. F., Hilsdorf, A. W. S., Liu, Z., & Coutinho, L. L. (2017). Large-scale SNP discovery and construction of a high-density genetic map of *Colossoma macropomum* through genotyping-by-sequencing. *Scientific Reports*, *7*. https://doi.org/10.1038/srep46112

Palumbi, S. R. (1994). *Genetic Divergence, Reproductive Isolation, and Marine Speciation*. 28.

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, *48*(1), 94–100. https://doi.org/10.1038/ng.3464

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Reboita, M. S., Krusche, N., Ambrizzi, T., & Rocha, R. P. da. (2015). Entendendo o tempo e o clima na América do Sul. *Terrae Didatica*, *8*(1), 34. https://doi.org/10.20396/td.v8i1.8637425

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Runnels, C. (2013). Phylogeography and Species Status of *Ramphogordius sanguineus*. *Theses and Dissertations*. https://doi.org/10.25772/EXFP-RG83

Strand, M., Norenburg, J., Alfaya, J. E., Ángel Fernández-Álvarez, F., Andersson, H. S., Andrade, S. C. S., Bartolomaeus, T., Beckers, P., Bigatti, G., Cherneva, I., Chernyshev, A., Chung, B. M., von Döhren, J., Giribet, G., Gonzalez-Cueto, J., Herrera-Bachiller, A., Hiebert, T., Hookabe, N., Junoy, J., … Xu, C.-M. (2019). Nemertean taxonomy-Implementing changes in the higher ranks, dismissing Anopla and Enopla. *Zoologica Scripta*, *48*(1), 118–119. https://doi.org/10.1111/zsc.12317

Thiel, M., & Junoy, J. (2006). Mating behavior of nemerteans: Present knowledge and future directions. *Journal of Natural History*, *40*(15–16), 1021–1034. https://doi.org/10.1080/00222930600834154

Thiel, M., & Kruse, I. (2001). Status of the Nemertea as predators in marine ecosystems. *Hydrobiologia*, *5*, 12.

Thollesson, M., & Norenburg, J. L. (2003). Ribbon worm relationships: A phylogeny of the phylum Nemertea. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1513), 407–415. https://doi.org/10.1098/rspb.2002.2254

Tigano, A., Shultz, A. J., Edwards, S. V., Robertson, G. J., & Friesen, V. L. (2017). Outlier analyses to test for local adaptation to breeding grounds in a migratory arctic seabird. *Ecology and Evolution*, *7*(7), 2370–2381. https://doi.org/10.1002/ece3.2819

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.2307/2408641

Yi, X., & Latch, E. K. (2021). Nonrandom missing data can bias Principal Component Analysis inference of population genetic structure. *Molecular Ecology Resources*, *00*(n/a), 1–10. https://doi.org/10.1111/1755-0998.13498

Zattara, E. E., Fernández-Álvarez, F. A., Hiebert, T. C., Bely, A. E., & Norenburg, J. L. (2019). A phylum-wide survey reveals multiple independent gains of head regeneration in Nemertea. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1898), 20182524. https://doi.org/10.1098/rspb.2018.2524

Zhbannikov, I. Y., Hunter, S. S., Foster, J. A., & Settles, M. L. (2017). SeqyClean: A Pipeline for High-throughput Sequence Data Preprocessing. *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics*, 407–416. https://doi.org/10.1145/3107411.3107446

SUPPLEMENTAL FILES
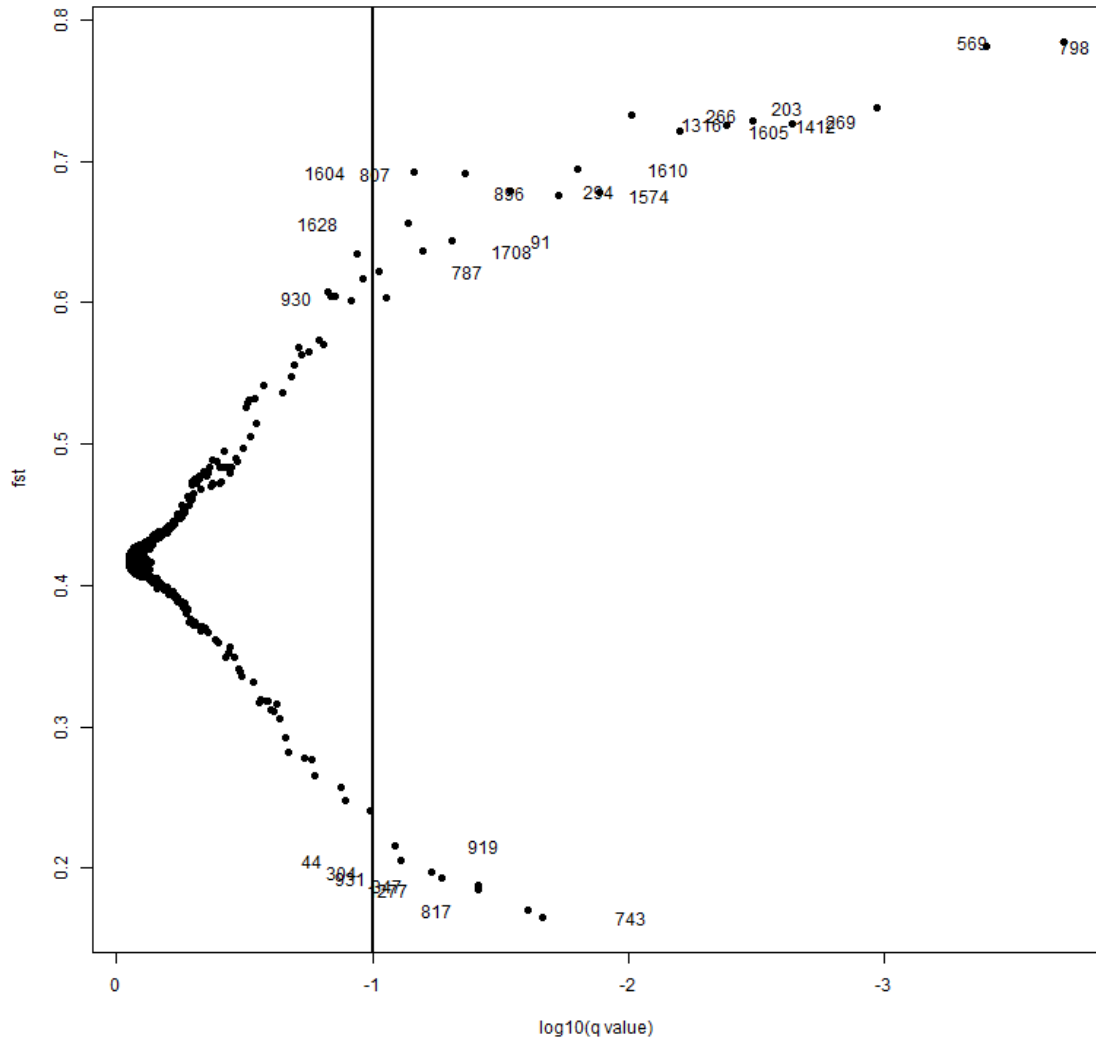
Figure S1. BayeScan plot of 1758 SNPs from 53 specimens of *Lineus sanguineus*. F_{ST} is plotted against the log10 of the False Discovery Rate (FDR). Vertical line shows the critical FDR used for outlier identification (FDR = 0.1).
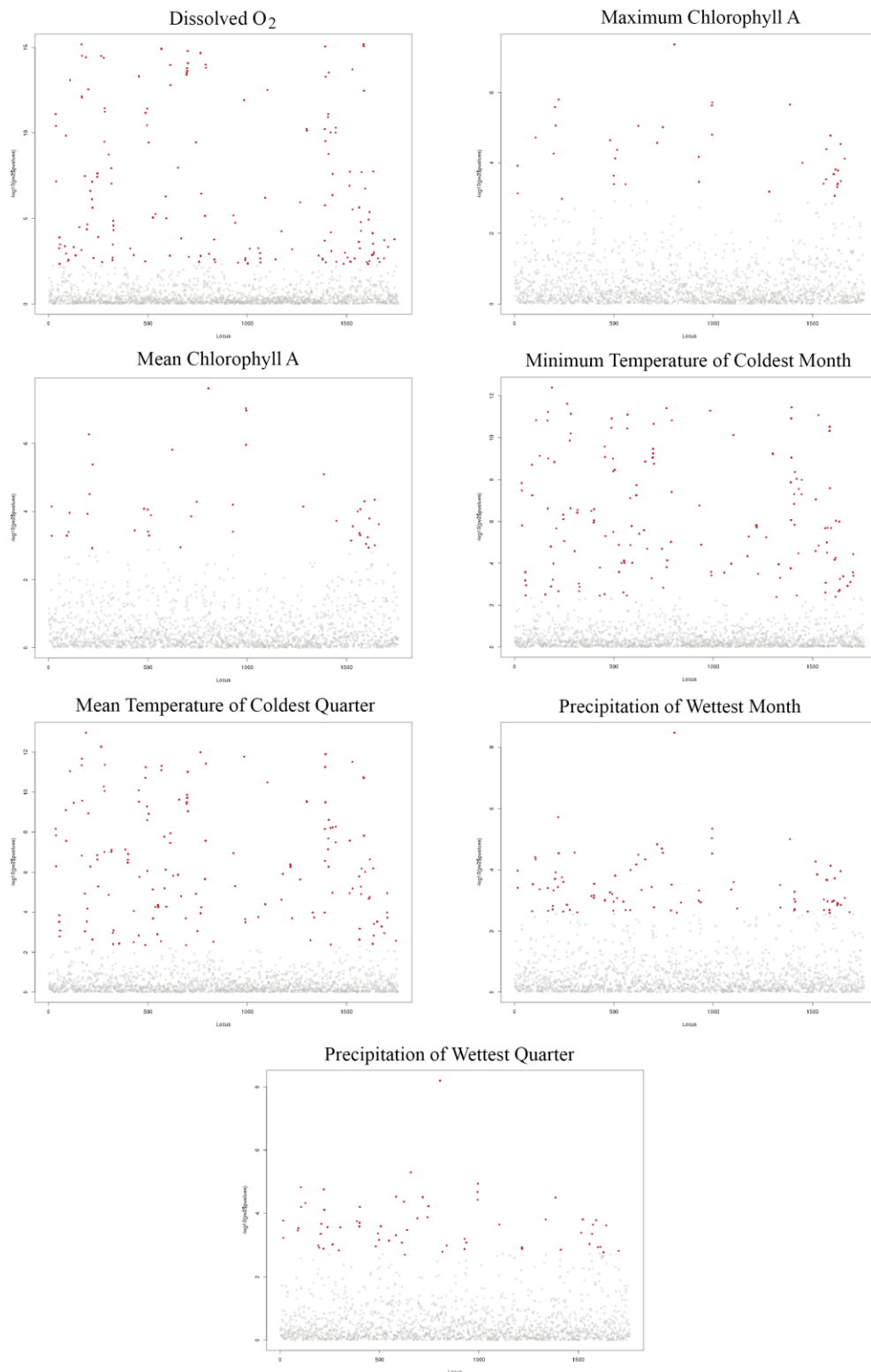
Figure S2. Manhattan plots of LFMM results using the three Bio-Oracle and four WorldClim variables described in table 5 and the 1758 SNPs from 53 specimens of *Lineus sanguineus*. Red dots represent outlier SNPs.
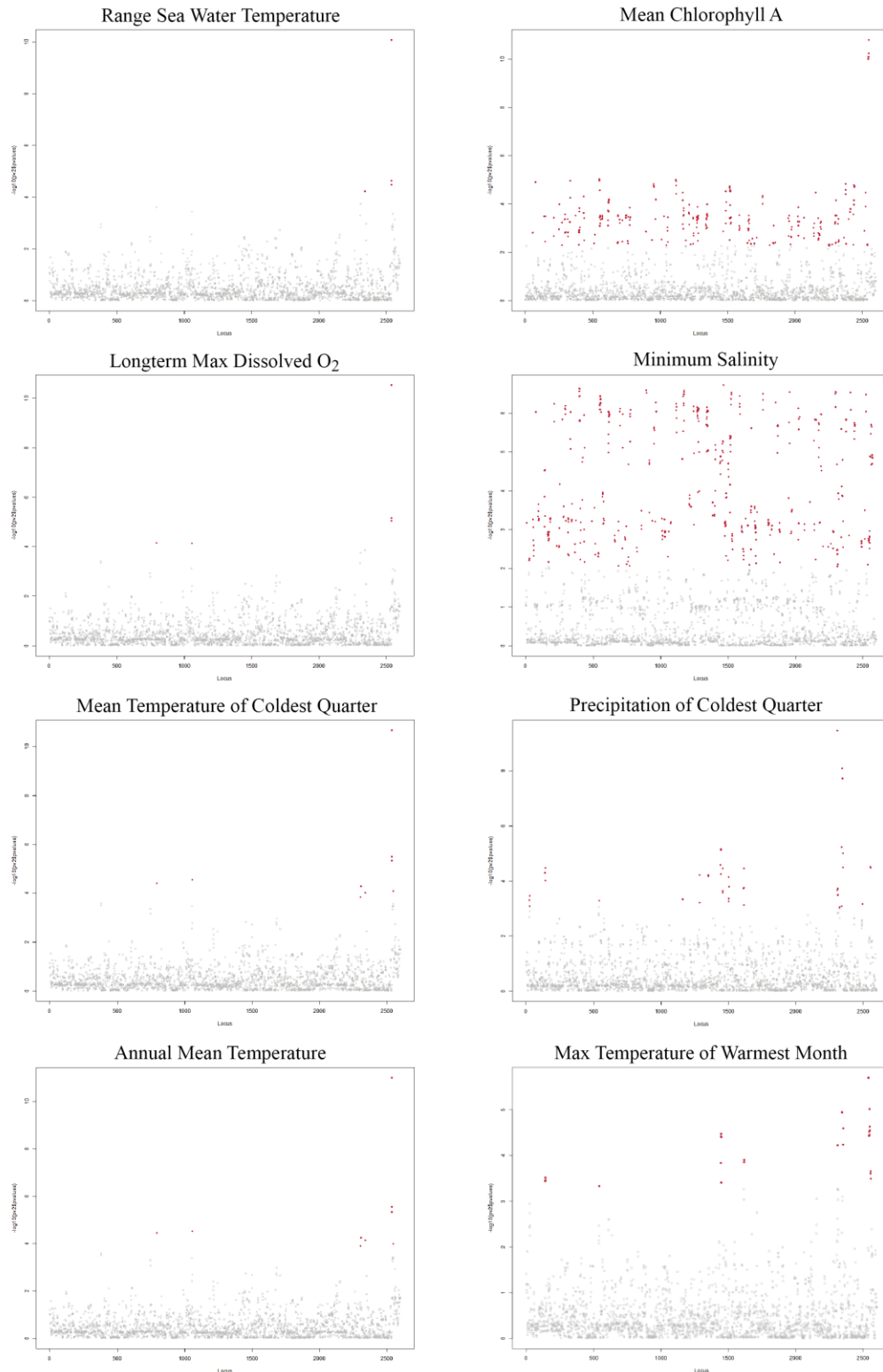
Figure S3. Manhattan plots of LFMM results using the four Bio-Oracle and four WorldClim variables described in table 5 and the 2601 SNPs from 19 specimens of *Nemertopsis berthalutzae*. Red dots represent outlier SNPs.
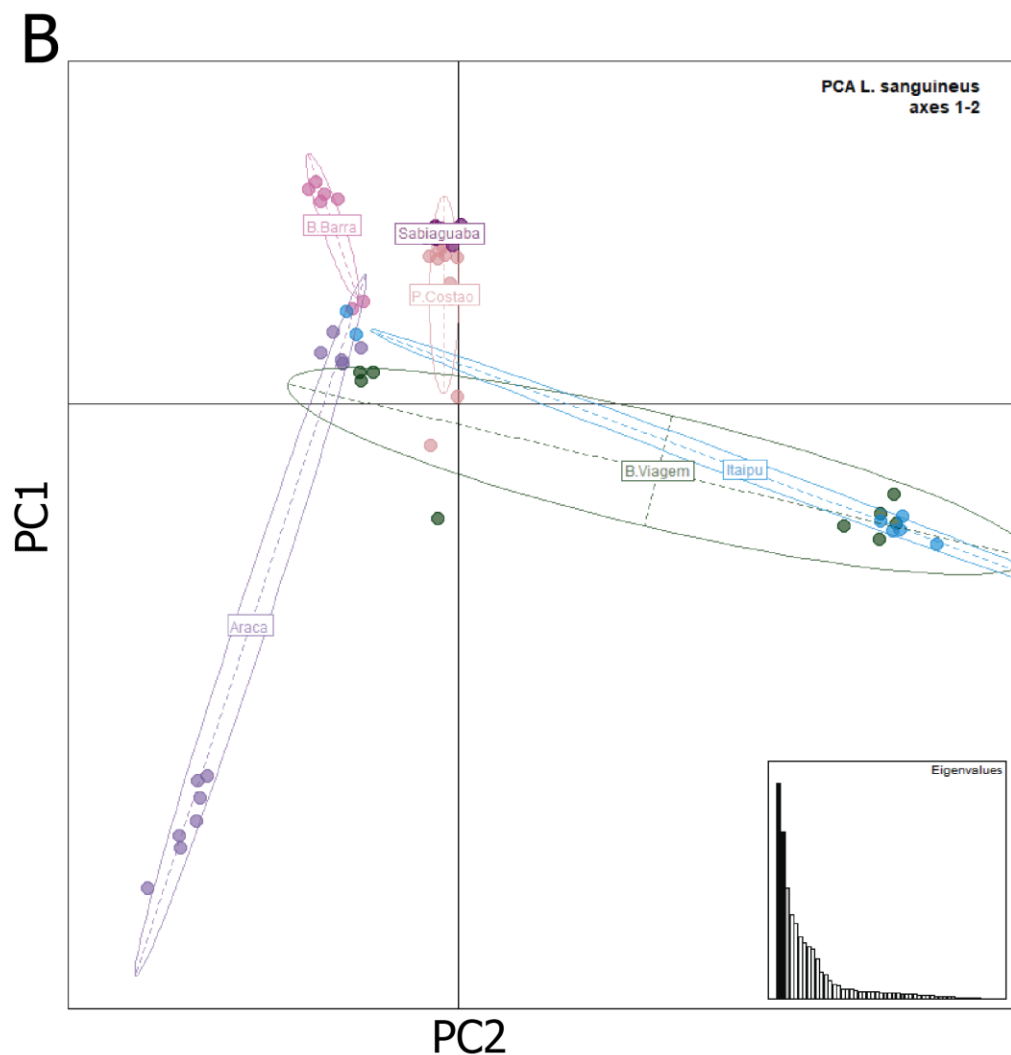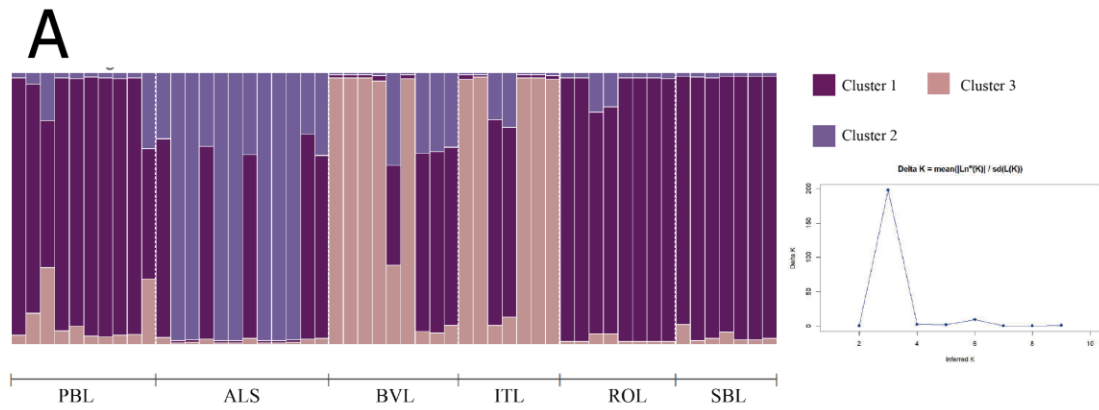
Figure S4. Clustering results for 274 outlier SNPs from *L. sanguineus* individuals. A: Structure results showing K=3 (left side) and ΔK results (right side); D: PCA results. Each circle corresponds to one individual, different colors represent different sample sites.
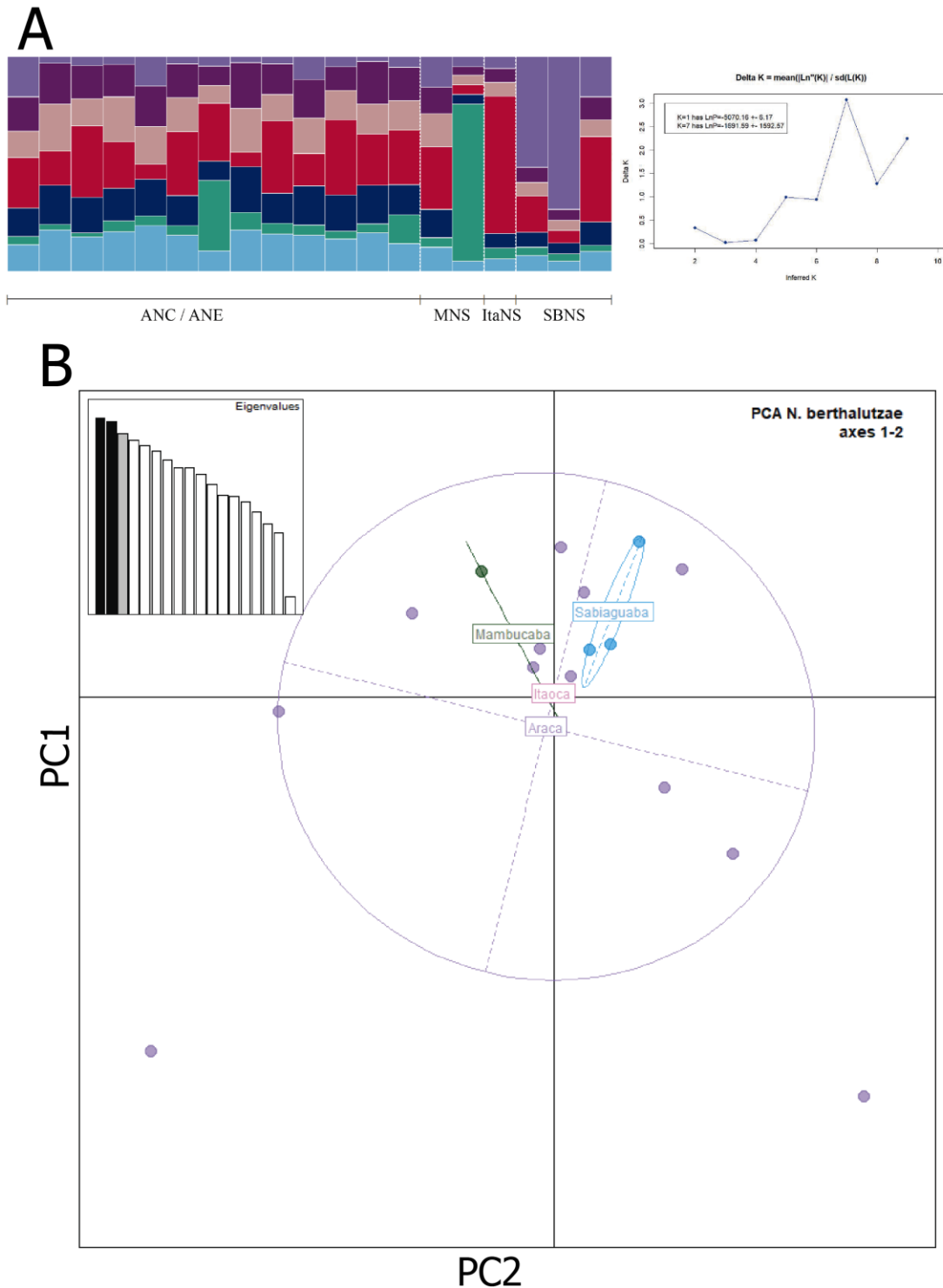
Figure S5. Clustering results for 519 outlier SNPs from *N. berthalutzae* individuals. A: Structure results showing K=7 (left side) and ΔK results, note insert with likelihoods for K=1 and K=7 (right side) B: PCA results. Each circle corresponds to one individual, different colors represent different sample sites.

Table S1. Number of sequenced reads per specimen before and after filtering steps, in the final consensus and final number of loci. Abbreviations per specimen as in table 1

.	Reads raw	Reads passed filter	Clusters total	Clusters hidepth	Reads consens	Loci in assembly
			Lineus sanguineus			
ALS1	2698743	2698723	66530	22383	21895	2043
ALS2	3383237	3383213	83643	24339	23723	2116
ALS3	7010816	7010760	97340	29737	29005	2126
ALS4	3229007	3228977	71754	24895	24377	2077
ALS5	1268311	1268303	67532	18604	18137	2120
ALS6	2493352	2493333	82005	22379	21799	2129
ALS7	1626311	1626302	67958	20009	19530	2135
ALS8	2116800	2116781	85949	22652	22125	2029
ALS9	1314855	1314839	60968	19526	19086	2042
ALS10	2377714	2377689	65422	22364	21852	2120
ALS11	4450345	4450311	67834	25432	24816	2131
ALS12	5606898	5606844	76390	26923	26236	2038
BVL1	2264459	2264346	26737	11506	11172	2141
BVL2	3425890	3425686	34581	13815	13381	2232
BVL3	2363980	2363835	33177	12747	12369	2219
BVL4	5171079	5170418	34733	15919	15467	1993
BVL5	2350001	2349816	33124	11673	11322	2081
BVL6	316950	316927	23805	6406	6157	1135
BVL7	6305622	6305297	35426	16461	16031	2150
BVL8	3895657	3895476	33973	13608	13248	2118
BVL9	1657219	1657138	32674	10552	10264	2062

BVL10	1823262	1823141	27408	10986	10675	2125
ITL3	3172917	3172689	55504	13782	13363	2202
ITL4	1134998	1134938	28661	8695	8482	1829
ITL5	4097252	4096999	32798	14810	14312	2234
ITL6	2494373	2494214	36512	13007	12622	2185
ITL8	6987013	6986587	64518	16703	16209	2226
ITL9	3228221	3228018	67298	14440	14045	2222
ITL10	2521639	2521492	54316	13588	13187	2212
PBL1	5212660	5212350	71183	16315	15821	2264
PBL2	1047068	1047005	28767	8639	8407	1694
PBL3	999656	999597	30869	9115	8855	1906
PBL4	2702754	2702609	29606	12852	12466	2254
PBL5	3459809	3459628	44070	15536	15120	2262
PBL6	3842026	3841802	61804	14097	13663	2244
PBL7	4900166	4899923	55131	16692	16225	2273
PBL8	6373263	6372931	44859	16459	15941	2254
PBL9	2485159	2485008	40175	11571	11199	2025
PBL10	1038475	1038400	29438	9732	9436	1950
ROL1	1246329	1246257	28916	8639	8405	1719
ROL2	675658	675621	21985	6358	6200	1360
ROL3	1283068	1282990	30342	8387	8168	1773
ROL4	1612975	1612893	27486	8501	8294	1720
ROL5	1438895	1438818	24130	8335	8132	1709
ROL6	1651362	1651290	23039	8455	8247	1839
ROL7	2039600	2039505	26529	9509	9298	1953
ROL8	2605323	2605189	29249	11022	10745	2060

ROL9	1010420	1010365	27172	7597	7412	1637
SBL1	4059316	4059105	49077	13469	13080	2192
SBL3	2489299	2489177	32593	12544	12217	2103
SBL4	2268691	2268541	31326	10921	10626	2127
SBL5	2569111	2568966	29854	10810	10505	2100
SBL6	2403176	2403035	35573	11595	11281	2114
SBL7	4526447	4526173	37867	13904	13524	2156
SBL8	4979321	4978945	46723	15900	15457	2166

Nemertopsis berthalutzae

.	Reads raw	Reads passed filter	Clusters total	Clusters hidepth	Reads consens	Loci in assembly
ANCI	3563793	3563224	18582	8255	7987	1146
ANCII	2539908	2539118	18475	7753	7586	1067
ANCIII	3117582	3117324	16553	6533	6361	1082
ANCIV	6353677	6352868	41306	12127	11765	1103
ANCV	6387024	6386416	89124	27612	27079	1084
ANCVI	8591471	8589146	23197	8228	7933	1090
ANCVII	1516035	1515828	15756	7304	7129	724
ANEI	2395800	2395644	24798	7453	7311	1055
ANEIII	5909122	5908220	73165	24747	24159	1149
ANEIV	3632801	3631951	46422	14799	14515	1087
ANEV	5583340	5582546	31954	14311	13977	1097
ANEVI	3639293	3638712	36118	16218	15901	1076
ANEVIII	7436830	7434559	45502	20095	19671	1081
ItaNS	625322	625222	14721	3417	3338	686
MNSI	4987192	4986920	43196	12269	12007	985

	Reads raw	Reads passed filter	Clusters total	Clusters hidepth	Reads consens	Loci in assembly
MNSII	1790799	1790548	39806	9220	9016	970
SBNSII	3005087	3003822	15213	5688	5513	1070
SBNSIII	4444532	4443599	34573	11474	11211	1074
SBNS4	1830908	1830630	14574	4237	4105	734

Nemertopsis pamelaroeae

.	Reads raw	Reads passed filter	Clusters total	Clusters hidepth	Reads consens	Loci in assembly
PNJ2	820537	820467	15392	2901	2839	605
ItaNJ	556441	556396	23017	7059	6956	635
GBN	542132	542104	16739	3897	3836	749
JNJI	820942	820898	21206	4491	4458	793
JNJIV	1604627	1604397	13472	4869	4805	947
JNJVI	2015856	2015759	25817	6128	6024	1128
JNJV	8988290	8987595	30136	9100	8946	1152
INJIV	7191822	7191424	44569	21315	20977	1166
JNJVII	3846306	3846136	26243	6436	6314	1196
JNJVIII	5786515	5786301	28652	8499	8347	1203
JNJIII	4087749	4087456	37201	9648	9474	1210
INJIII	3365552	3365367	51299	15462	15045	1211
INJI	2132986	2132817	33083	11012	10866	1227
INJII	4511961	4511697	76469	21996	21555	1242

Table S2. Environmental variables used in Principal Component Analyses of *Lineus sanguineus*, *Nemertopsis berthalutzae* and *Nemertopsis pamelaroeae* distribution from Bio-Oracle and World Clim databases.

Bio-Oracle Environmental Variables	
Layer Code	**Variable Name**
BO_chlomax	Chlorophyll A (maximum)
BO_chlomean	Chlorophyll A (mean)
BO_chlomin	Chlorophyll A (minimum)
BO_chlorange	Chlorophyll A (range)
BO_nitrate	Nitrate
BO_ph	pH
BO_sstrange	Sea surface temperature (range)
BO2_carbonphytomax_bdmin	Carbon phytoplankton biomass (maximum at min depth)
BO2_carbonphytomean_bdmin	Carbon phytoplankton biomass (mean at mean depth)
BO2_carbonphytorange_bdmin	Carbon phytoplankton biomass (range at min depth)
BO2_chlomin_bdmin	Chlorophyll concentration (minimum at min depth)
BO2_chlomax_bdmin	Chlorophyll concentration (maximum at min depth)
BO2_chlorange_bdmin	Chlorophyll concentration (range at min depth)
BO2_dissoxltmax_bdmin	Dissolved oxygen concentration (longterm max at min depth)
BO2_dissoxltmin_bdmin	Dissolved oxygen concentration (longterm min at min depth)
BO2_dissoxmax_bdmin	Dissolved oxygen concentration (maximum at min depth)
BO2_dissoxmean_bdmin	Dissolved oxygen concentration (mean at min depth)
BO2_dissoxmin_bdmin	Dissolved oxygen concentration (minimum at min depth)
BO2_dissoxrange_bdmin	Dissolved oxygen concentration (range at min depth)
BO2_nitratemax_bdmin	Nitrate concentration (maximum at min depth)
BO2_nitraterange_bdmin	Nitrate concentration (range at min depth)
BO2_salinitymean_ss	Sea surface salinity (mean)
BO2_temprange_bdmin	Sea water temperature (range at min depth)
BO2_temprange_ss	BO2_temprange_ss

World Clim Biological variables	
Layer Code	**Environmental Variable**
bio1	Annual Mean Temperature
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
bio3	Isothermality (BIO2/BIO7) ($\times100$)
bio4	Temperature Seasonality (standard deviation $\times100$)
bio5	Max Temperature of Warmest Month
bio6	Min Temperature of Coldest Month
bio7	Temperature Annual Range (BIO5-BIO6)
bio8	Mean Temperature of Wettest Quarter
bio9	Mean Temperature of Driest Quarter
bio10	Mean Temperature of Warmest Quarter
bio11	Mean Temperature of Coldest Quarter
bio12	Annual Precipitation
bio13	Precipitation of Wettest Month
bio14	Precipitation of Driest Month
bio15	Precipitation Seasonality (Coefficient of Variation)
bio16	Precipitation of Wettest Quarter
bio17	Precipitation of Driest Quarter
bio18	Precipitation of Warmest Quarter
bio19	Precipitation of Coldest Quarter

Table S3. Blast hits of the outlier SNPs identified through Blast searches against NT and NR databases, and in house transcriptomes of *Lineus sanguineus*.

Source of outlier SNP	Blast Hit	UniProt Info
Dissolved Oxygen	Ankyrin-2;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	axon extension / microtubule cytoskeleton organization / signal transduction
Dissolved Oxygen	Transient receptor potential cation channel subfamily M member-like 2 {ECO:0000305|PubMed:25620041};^Eukaryota; Metazoa; Cnidaria; Anthozoa; Hexacorallia; Actiniaria; Edwardsiidae; Nematostella	Functions as ligand-gated ion channel
Dissolved Oxygen	Lysine-specific demethylase 5A;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo`KDM5A_HUMAN^KDM5A_HUMAN^Q:4899-5084,H:1599-1662^44.62%ID^E:5e-09^RecName: Full=Lysine-specific demethylase 5A;	May stimulate transcription mediated by nuclear receptors. Involved in transcriptional regulation of Hox proteins during cell differentiation / Regulates specific gene transcription through DNA-binding on 5'-CCGCCC-3' motif
Dissolved Oxygen	Papilin;^Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Holometabola; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila; Sophophora	extracellular matrix organization
Dissolved Oxygen	CD9 antigen {ECO:0000303|PubMed:1339429};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	In myoblasts, associates with CD81 and PTGFRN and inhibits myotube fusion during muscle regeneration / Involved in cell adhesion, cell motility and tumor metastasis

Dissolved Oxygen	Transcription initiation factor TFIID subunit 4;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	TFIID is a multimeric protein complex that plays a central role in mediating promoter responses to various activators and repressors. May function as a coactivator by serving as a site of protein-protein contact between activators like Sp1 (or btd) and TFIID complex
Dissolved Oxygen	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Protein of the nuclear speckles that regulates RNA degradation and export from the nucleus through its interaction with MTREX an essential factor directing various RNAs to exosomal degradation / Plays a role in DNA damage response
Dissolved Oxygen	Transient receptor potential cation channel subfamily M member-like 2 {ECO:0000305|PubMed:25620041};^Eukaryota; Metazoa; Cnidaria; Anthozoa; Hexacorallia; Actiniaria; Edwardsiidae; Nematostella	May have ADP-ribose pyrophosphatase activity which reduces ADP-ribose levels induced by oxidative stress, thus preventing the channel activation by reactive oxygen species
Dissolved Oxygen	Citron Rho-interacting kinase;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	Plays a role in cytokinesis.

Dissolved Oxygen	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Protein of the nuclear speckles that regulates RNA degradation and export from the nucleus through its interaction with MTREX an essential factor directing various RNAs to exosomal degradation / Plays a role in DNA damage response
Dissolved Oxygen	Tolloid-like protein 2;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	cell differentiation
Dissolved Oxygen	Synaptopodin 2-like protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Actin-associated protein that may play a role in modulating actin-based shape
Dissolved Oxygen	Probable ubiquitin carboxyl-terminal hydrolase FAF-X;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	May therefore play an important regulatory role at the level of protein turnover by preventing degradation of proteins through the removal of conjugated ubiquitin. / thereby promoting the repair of alkylated DNA lesions

Dissolved Oxygen	YEATS domain-containing protein 2 {ECO:0000305};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	negative regulation of transcription, DNA-templated / Chromatin reader component of the ATAC complex, a complex with histone acetyltransferase activity on histones H3 and H4
Dissolved Oxygen	G2/mitotic-specific cyclin-B3;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves; Neognathae; Galloanserae; Galliformes; Phasianidae; Phasianinae; Gallus	Cyclins are positive regulatory subunits of the cyclin-dependent kinases (CDKs), and thereby play an essential role in the control of the cell cycle, notably via their destruction during cell division. Probably functions redundantly with other cyclins in regulation of cell cycle. Its presence may be required to delay a deadline for completing cytokinesis that is ordinary imposed by nuclear envelope reformation. Degradation of CycB and CycB3 promote cytokinesis furrow initiation and ingression

Dissolved Oxygen	Cancer-related nucleoside-triphosphatase homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	Has nucleotide phosphatase activity towards ATP, GTP, CTP, TTP and UTP. Hydrolyzes nucleoside diphosphates with lower efficiency
Dissolved Oxygen	DNA polymerase alpha catalytic subunit;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Catalytic subunit of the DNA polymerase alpha complex (also known as the alpha DNA polymerase-primase complex) which plays an essential role in the initiation of DNA synthesis / The primase subunit of the polymerase alpha complex initiates DNA synthesis by oligomerising short RNA primers on both leading and lagging strands
Dissolved Oxygen	B-cell linker protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Plays a critical role in orchestrating the pro-B cell to pre-B cell transition. May play an important role in BCR-induced B-cell apoptosis

Dissolved Oxygen	Neuroendocrine convertase 2 {ECO:0000250|UniProtKB:P16519};^Eukaryota; Metazoa; Ecdysozoa; Nematoda; Chromadorea; Rhabditida; Rhabditina; Rhabditomorpha; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis	Serine endoprotease which cleaves preproteins at paired basic amino acids / Probably by processing flp-1 and flp-18, modulates the neuronal excitation-inhibition balance and thus the level of activity of the locomotor circuit / Regulates sensitivity to mechanosensory stimuli
Dissolved Oxygen	Nuclear transcription factor Y subunit alpha;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	NF-Y can function as both an activator and a repressor, depending on its interacting cofactors. NF-YA positively regulates the transcription of the core clock component ARNTL/BMAL1. / ciclo cicardiano
Maximum Clorophyll A	snRNA-activating protein complex subunit 4;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Part of the SNAPc complex required for the transcription of both RNA polymerase II and III small-nuclear RNA genes. Binds to the proximal sequence element (PSE), a non-TATA-box basal promoter element common to these 2 types of genes. Recruits TBP and BRF2 to the U6 snRNA TATA box

Maximum Clorophyll A	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Maximum Clorophyll A	Corticotropin-releasing factor-binding protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Binds CRF and inactivates it. / Binds CRF and inactivates it. / hormone metabolic process / regulation of cellular response to stress
Maximum Clorophyll A	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Maximum Clorophyll A	Collagen alpha-6(VI) chain;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	Collagen VI acts as a cell-binding protein.
Maximum Clorophyll A	Echinoderm microtubule-associated protein-like 6;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	May modify the assembly dynamics of microtubules, such that microtubules are slightly longer, but more dynamic.

Mean Clorophyll A	snRNA-activating protein complex subunit 4;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Part of the SNAPc complex required for the transcription of both RNA polymerase II and III small-nuclear RNA genes. Binds to the proximal sequence element (PSE), a non-TATA-box basal promoter element common to these 2 types of genes. Recruits TBP and BRF2 to the U6 snRNA TATA box.
Minimum Temperature of Coldest Month	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Minimum Temperature of Coldest Month	Corticotropin-releasing factor-binding protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Binds CRF and inactivates it. / Binds CRF and inactivates it. / hormone metabolic process / regulation of cellular response to stress
Minimum Temperature of Coldest Month	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Minimum Temperature of Coldest Month	Collagen alpha-6(VI) chain;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	Collagen VI acts as a cell-binding protein.

Minimum Temperature of Coldest Month	Echinoderm microtubule-associated protein-like 6;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	May modify the assembly dynamics of microtubules, such that microtubules are slightly longer, but more dynamic.
Minimum Temperature of Coldest Month	Lysine-specific demethylase 5A;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Regulates specific gene transcription through DNA-binding on 5'-CCGCCC-3' motif / Plays a role in the regulation of the circadian rhythm and in maintaining the normal periodicity of the circadian clock. In a histone demethylase-independent manner, acts as a coactivator of the CLOCK-ARNTL/BMAL1-mediated transcriptional activation of PER1/2 and other clock-controlled genes and increases histone acetylation at PER1/2 promoters by inhibiting the activity of HDAC1
Minimum Temperature of Coldest Month	Papilin;^Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Holometabola; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila; Sophophora	extracellular matrix organization

Minimum Temperature of Coldest Month	CD9 antigen {ECO:0000303|PubMed:1339429};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	In myoblasts, associates with CD81 and PTGFRN and inhibits myotube fusion during muscle regeneration / Involved in cell adhesion, cell motility and tumor metastasis
Minimum Temperature of Coldest Month	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Protein of the nuclear speckles that regulates RNA degradation and export from the nucleus through its interaction with MTREX an essential factor directing various RNAs to exosomal degradation / Plays a role in DNA damage response

Minimum Temperature of Coldest Month	Citron Rho-interacting kinase;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	Plays a role in cytokinesis. Required for KIF14 localization to the central spindle and midbody. Probable RHO/RAC effector that binds to the GTP-bound forms of RHO and RAC1. It probably binds p21 with a tighter specificity in vivo. Displays serine/threonine protein kinase activity. Plays an important role in the regulation of cytokinesis and the development of the central nervous system. Phosphorylates MYL9/MLC2.
Minimum Temperature of Coldest Month	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	cell division / cellular response to DNA damage stimulus / Changes the conformation of MTREX, precluding its association with the nuclear exosome and interaction with proteins required for its function in RNA exosomal degradation
Minimum Temperature of Coldest Month	Tolloid-like protein 2;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	cell differentiation

| Minimum Temperature of Coldest Month | Probable ubiquitin carboxyl-terminal hydrolase FAF-X;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo | Amyloid fibril formation / axon extension / BMP signaling pathway / cell cycle / cell division / cell migration / chromosome segregation / May therefore play an important regulatory role at the level of protein turnover by preventing degradation of proteins through the removal of conjugated ubiquitin / Deubiquitinates alkylation repair enzyme ALKBH3. OTUD4 recruits USP7 and USP9X to stabilize ALKBH3, thereby promoting the repair of alkylated DNA lesions |
| Minimum Temperature of Coldest Month | RING finger protein 10;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus | Transcriptional factor involved in the regulation of MAG (Myelin-associated glycoprotein) expression. Acts as a regulator of Schwann cell differentiation and myelination. / positive regulation of transcription, DNA-templated |

Minimum Temperature of Coldest Month	YEATS domain-containing protein 2 {ECO:0000305};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Chromatin reader component of the ATAC complex, a complex with histone acetyltransferase activity on histones H3 and H4 / negative regulation of transcription, DNA-templated
Minimum Temperature of Coldest Month	G2/mitotic-specific cyclin-B3;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves; Neognathae; Galloanserae; Galliformes; Phasianidae; Phasianinae; Gallus	Could be involved at the G2/M (mitosis) transition (Probable). Interacts with the CDK1 and CDK2 protein kinases (Probable). G2/M cyclins accumulate steadily during G2 and are abruptly destroyed at mitosis (Probable). Plays a role during oocyte meiosis II
Minimum Temperature of Coldest Month	Cancer-related nucleoside-triphosphatase homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	Has nucleotide phosphatase activity towards ATP, GTP, CTP, TTP and UTP. Hydrolyzes nucleoside diphosphates with lower efficiency
Minimum Temperature of Coldest Month	DNA polymerase alpha catalytic subunit;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Plays an essential role in the initiation of DNA replication. /

Minimum Temperature of Coldest Month	B-cell linker protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Functions as a central linker protein, downstream of the B-cell receptor (BCR), bridging the SYK kinase to a multitude of signaling pathways and regulating biological outcomes of B-cell function and development.
Minimum Temperature of Coldest Month	Neuroendocrine convertase 2 {ECO:0000250|UniProtKB:P16519};^Eukaryota; Metazoa; Ecdysozoa; Nematoda; Chromadorea; Rhabditida; Rhabditina; Rhabditomorpha; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis	Serine endoprotease which cleaves preproteins at paired basic amino acids / Probably by processing flp-1 and flp-18, modulates the neuronal excitation-inhibition balance and thus the level of activity of the locomotor circuit / Regulates sensitivity to mechanosensory stimuli /
Minimum Temperature of Coldest Month	Nuclear transcription factor Y subunit alpha;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	NF-Y can function as both an activator and a repressor, depending on its interacting cofactors. NF-YA positively regulates the transcription of the core clock component ARNTL/BMAL1.

Minimum Temperature of Coldest Month	Nuclear transcription factor Y subunit alpha;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	NF-Y can function as both an activator and a repressor, depending on its interacting cofactors. NF-YA positively regulates the transcription of the core clock component ARNTL/BMAL1.
Minimum Temperature of Coldest Month	Probable ATP-dependent RNA helicase DHX34;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	negative regulation of nuclear-transcribed mRNA catabolic process, nonsense-mediated decay

Mean Temperature of Coldest Quarter	Lysine-specific demethylase 5A;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Regulates specific gene transcription through DNA-binding on 5'-CCGCCC-3' motif / Plays a role in the regulation of the circadian rhythm and in maintaining the normal periodicity of the circadian clock. In a histone demethylase-independent manner, acts as a coactivator of the CLOCK-ARNTL/BMAL1-mediated transcriptional activation of PER1/2 and other clock-controlled genes and increases histone acetylation at PER1/2 promoters by inhibiting the activity of HDAC1
Mean Temperature of Coldest Quarter	Papilin;^Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Holometabola; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila; Sophophora	extracellular matrix organization
Mean Temperature of Coldest Quarter	CD9 antigen {ECO:0000303|PubMed:1339429};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	In myoblasts, associates with CD81 and PTGFRN and inhibits myotube fusion during muscle regeneration / Involved in cell adhesion, cell motility and tumor metastasis

Mean Temperature of Coldest Quarter	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Protein of the nuclear speckles that regulates RNA degradation and export from the nucleus through its interaction with MTREX an essential factor directing various RNAs to exosomal degradation / Plays a role in DNA damage response
Mean Temperature of Coldest Quarter	Citron Rho-interacting kinase;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	Plays a role in cytokinesis. Required for KIF14 localization to the central spindle and midbody. Probable RHO/RAC effector that binds to the GTP-bound forms of RHO and RAC1. It probably binds p21 with a tighter specificity in vivo. Displays serine/threonine protein kinase activity. Plays an important role in the regulation of cytokinesis and the development of the central nervous system. Phosphorylates MYL9/MLC2.

Mean Temperature of Coldest Quarter	Protein NRDE2 homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	cell division / cellular response to DNA damage stimulus / Changes the conformation of MTREX, precluding its association with the nuclear exosome and interaction with proteins required for its function in RNA exosomal degradation
Mean Temperature of Coldest Quarter	Tolloid-like protein 2;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	cell differentiation
Mean Temperature of Coldest Quarter	Hepatoma-derived growth factor-related protein 2;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana	regulation of transcription by RNA polymerase II

Mean Temperature of Coldest Quarter	Probable ubiquitin carboxyl-terminal hydrolase FAF-X;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	amyloid fibril formation / axon extension / BMP signaling pathway / cell cycle / cell division / cell migration / chromosome segregation / May therefore play an important regulatory role at the level of protein turnover by preventing degradation of proteins through the removal of conjugated ubiquitin / Deubiquitinates alkylation repair enzyme ALKBH3. OTUD4 recruits USP7 and USP9X to stabilize ALKBH3, thereby promoting the repair of alkylated DNA lesions
Mean Temperature of Coldest Quarter	RING finger protein 10;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Transcriptional factor involved in the regulation of MAG (Myelin-associated glycoprotein) expression. Acts as a regulator of Schwann cell differentiation and myelination. / positive regulation of transcription, DNA-templated

Mean Temperature of Coldest Quarter	YEATS domain-containing protein 2 {ECO:0000305};^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Chromatin reader component of the ATAC complex, a complex with histone acetyltransferase activity on histones H3 and H4 / negative regulation of transcription, DNA-templated
Mean Temperature of Coldest Quarter	G2/mitotic-specific cyclin-B3;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda; Coelurosauria; Aves; Neognathae; Galloanserae; Galliformes; Phasianidae; Phasianinae; Gallus	play an essential role in the control of the cell cycle, notably via their destruction during cell division. / cell division
Mean Temperature of Coldest Quarter	Cancer-related nucleoside-triphosphatase homolog;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos	Has nucleotide phosphatase activity towards ATP, GTP, CTP, TTP and UTP. Hydrolyzes nucleoside diphosphates with lower efficiency
Mean Temperature of Coldest Quarter	DNA polymerase alpha catalytic subunit;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Catalytic subunit of the DNA polymerase alpha complex (also known as the alpha DNA polymerase-primase complex) which plays an essential role in the initiation of DNA synthesis / The primase subunit of the polymerase alpha complex initiates DNA synthesis by oligomerising short RNA primers on both leading and lagging strands

Mean Temperature of Coldest Quarter	B-cell linker protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	inflammatory response / B cell receptor signaling pathway / B cell differentiation
Mean Temperature of Coldest Quarter	Neuroendocrine convertase 2 {ECO:0000250|UniProtKB:P16519};^Eukaryota; Metazoa; Ecdysozoa; Nematoda; Chromadorea; Rhabditida; Rhabditina; Rhabditomorpha; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis	Serine endoprotease which cleaves preproteins at paired basic amino acids / Probably by processing flp-1 and flp-18, modulates the neuronal excitation-inhibition balance and thus the level of activity of the locomotor circuit / Regulates sensitivity to mechanosensory stimuli /
Mean Temperature of Coldest Quarter	Nuclear transcription factor Y subunit alpha;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	NF-Y can function as both an activator and a repressor, depending on its interacting cofactors. NF-YA positively regulates the transcription of the core clock component ARNTL/BMAL1.
Mean Temperature of Coldest Quarter	Probable ATP-dependent RNA helicase DHX34;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	negative regulation of nuclear-transcribed mRNA catabolic process, nonsense-mediated decay

Mean Temperature of Coldest Quarter	Rho GTPase-activating protein 20;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	signal transduction
Precipitation of Wettest Month	Coiled-coil domain-containing protein 22 homolog;^Eukaryota; Metazoa; Cnidaria; Anthozoa; Hexacorallia; Actiniaria; Edwardsiidae; Nematostella	positive regulation of ubiquitin-dependent protein catabolic process
Precipitation of Wettest Month	snRNA-activating protein complex subunit 4;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Binds to the proximal sequence element (PSE), a non-TATA-box basal promoter element common to these 2 types of genes. Recruits TBP and BRF2 to the U6 snRNA TATA box /
Precipitation of Wettest Month	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Precipitation of Wettest Month	Corticotropin-releasing factor-binding protein;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus	Binds CRF and inactivates it. / Binds CRF and inactivates it. / hormone metabolic process / regulation of cellular response to stress

Precipitation of Wettest Month	Citron Rho-interacting kinase;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus	Plays a role in cytokinesis. Required for KIF14 localization to the central spindle and midbody. Probable RHO/RAC effector that binds to the GTP-bound forms of RHO and RAC1. It probably binds p21 with a tighter specificity in vivo. Displays serine/threonine protein kinase activity. Plays an important role in the regulation of cytokinesis and the development of the central nervous system. Phosphorylates MYL9/MLC2.
Precipitation of Wettest Month	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Precipitation of Wettest Month	Collagen alpha-6(VI) chain;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	Collagen VI acts as a cell-binding protein.

Precipitation of Wettest Month	Metabotropic glutamate receptor;^Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Holometabola; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila; Sophophora	G-protein coupled receptor for glutamate. Ligand binding causes a conformation change that triggers signaling via guanine nucleotide-binding proteins (G proteins) and modulates the activity of down-stream effectors. / regulation of synaptic transmission, glutamatergic / short-term memory
Precipitation of Wettest Quarter	Coiled-coil domain-containing protein 22 homolog;^Eukaryota; Metazoa; Cnidaria; Anthozoa; Hexacorallia; Actiniaria; Edwardsiidae; Nematostella	positive regulation of ubiquitin-dependent protein catabolic process
Precipitation of Wettest Quarter	Sacsin;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo	Co-chaperone which acts as a regulator of the Hsp70 chaperone machinery and may be involved in the processing of other ataxia-linked proteins
Precipitation of Wettest Quarter	Collagen alpha-6(VI) chain;^Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus	Collagen VI acts as a cell-binding protein.

General Conclusions

In the present study we evaluated the validity of the three selected species (*Lineus sanguineus*, *Nemertopsis bivittata* and *Perinereis ponteni*), their populational structure along the Brazilian coast, as well as their past demography and possible responses to their environment.

Our first results show that *N. bivittata* is a complex of cryptic species comprehending three different species in Brazil, *N. berthalutzae*, *N. caete* and *N. pamelaroeae*. From those three species, only two seem to be present along the entire sampling range, *N. berthalutzae* and *N. pamelaroeae*, being *N. caete* restrict to the coast of Alagoas state.

The hoplonemertean larval development was not yet well known, except for a few species with direct development. However, many populational results, as well as some anecdotal observations of larval growth in the plankton, indicated indirect development. Therefore, we investigated the hoplonemertean development using *Emplectonema viride* as a model species. The observation of *E. viride* larval development indicated a long planktonic stage of around 120 days, with feeding starting as soon as the stylet developed, at four days old. These results, as well as the ones described in von Dassow *et al.* (in press) indicated that planktonic larval development is well spread among the hoplonemerteans. The long planktonic stage observed is compatible to the populational results found for other hoplonemertean species that have very well connected populations, despite the low motile adults.

The population results from all four species (*P. ponteni*, *L. sanguineus*, *N. berthalutzae* and *N. pamelaroeae*) reaffirm the importance of developmental modes to the population dynamics. Species with longer larval stage (i.e. the two

planktotrophic *Nemertopsis* species) have more connected populations than species with shorter larval stage (i.e. the two lecithotrophic *L. sanguineus* and *P. ponteni*), despite having similar habits during the adult life. Nonetheless, our results also shed light in other factors that also shape the population dynamics, particularly the past demography and the present environment.

The demography results for the *Nemertopsis* species are not informative, due to the uneven sampling. However, both our results for *P. ponteni* and *L. sanguineus* suggest a demographic expansion from the Southeast region, where the populations show higher levels of diversity. These levels of diversity are most likely consequence of this area being a refugee for the species during the Last Glacial Maximum (~ 20ky), when the Brazilian coast was more exposed than nowadays and the suitable habitats for coastal species were rarer.

Currently, the temperature seems to be main environmental factor that influence all studied species. They all show SNPs candidates to be under selection related directly or indirectly to temperature variation. The SNPs are related to many different genes in the different species. Nevertheless, genes related to DNA repair and immunity are present in all cases – *ZBRAN3* and *PARP9* among *P. ponteni* populations; *TMP2L*, *NRDE2*, *USP9X*, *SACS*, *CRHBP*, *CD9* and *BLNK* in *L. sanguineus*; *T184B*, *TENS1*, *MACF1* and *DYH7* amongst *N. berthalutzae* populations – suggesting a common response to the challenging environment in all studied species, despite their differences. The increase in temperature favors primary production and microorganism proliferation, which can increase the likelihood of disease in environments already highly impacted by human interventions. In addition, fouling communities provide a propitious microhabitat for accumulation of organic matter that also favors such microorganisms. Besides the local effect of temperature

increase, the changing in the ocean temperature can also have a long term effect in these populations, due to the alteration in the oceanic circulation. We cannot yet affirm how this alteration will affect the populations. However, the temperature difference is a very important factor to the direction and velocity of ocean currents, and a warmer ocean is likely to have slower currents, isolating the animal populations that use the present ocean circulation for connection. It is interesting to note that even a species with such a high gene flow still responds to biogeographic barriers, highlighting the importance of biotic and abiotic factors shaping population connectivity. It is also worth noting the importance of temperature in shaping these populations, indicating that future climate change and ocean warming can have huge impacts on coastal communities.

Resumo

Entender quais são os principais fatores que influenciam a conectividade entre populações animais marinhas tem sido objeto de diversos estudos há décadas, sendo o a mobilidade dos indivíduos reconhecida como um dos fatores determinantes. Espécies com longo desenvolvimento larval (e.g. planctotróficas) usualmente apresentam populações melhor conectadas que espécies com desenvolvimento direto ou curto desenvolvimento larval (e.g. lecitotróficas), devido sua maior capacidade de dispersão. Entretanto, outras características bionômicas e o ambiente também tem grande influência na conectividade. Além disso, a percepção da conectividade populacional também sofre interferência devido a presença de espécies crípticas, uma vez que tais espécies podem ter extrema semelhança morfológica. Assim, o presente trabalho avaliou como a heterogeneidade ambiental influencia a conectividade de espécies com diferentes modos reprodutivos, através de técnicas de genômica de paisagem comparativa. Para tanto, utilizamos as espécies *Perinereis ponteni* (Polychaeta, de reprodução sexuada e larva lecitotrófica), *Lineus sanguineus* (Heteronemertea, de reprodução assexuada e sexuada, e larva lecitotrófica) e *Nemertopsis bivittata* (Hoplonemertea, de reprodução sexuada, com desenvolvimento larval não completamente conhecido, mas presumidamente planctotrófica). De forma a garantir a monofilia das espécies, as populações estudadas foram avaliadas taxonomicamente. Nesse sentido, amostramos populações ao longo da costa brasileira. Os animais coletados foram utilizados na extração de DNA para prospecção de SNPs, a partir da técnica de *Genotyping by Sequencing* (GBS). Ainda, acompanhamos o comportamento e desenvolvimento larval do hoplonemertíneo *Emplectonema viride* como proxy para o desenvolvimento de *N. bivittata*, uma vez que ambas são da família Emplectonematidae e vivem em ambientes extremamente similares. Contudo, a espécie *N. bivittata* é na verdade um complexo de espécies, com três unidades evolutivas distintas na costa brasileira. Essas foram descritas como, *Nemertopsis berthalutzae*, *Nemertopsis pamelaroeae* e *Nemertopsis caete*, a última estando presente apenas na costa de Alagoas. A observação do desenvolvimento larval de *E. viride* indicou a presença de larvas planctotróficas, com longo desenvolvimento, levando em média 120 dias para metamorfose bastante sutil (notada principalmente pela mudança na ciliação e no comportamento). As larvas se alimentam primariamente de nauplius e ciprídes de cracas. Nossas análises de paleodistribuição e demografia de *P. ponteni* apontaram a presença de um refúgio na região Sudeste durante o Último Máximo Glacial, de onde as populações expandiram. Quanto às análises de genômica de paisagem, observamos que as espécies com desenvolvimento larval mais curto tendem a apresentar populações mais estruturadas, porém ainda com significativo fluxo gênico. Entretanto, fatores ambientais semelhantes afetam a conectividade em todas as espécies, sendo a temperatura e precipitação os mais comuns. Esses fatores estão associados à presença de SNPs candidatos à seleção natural em três das quatro espécies estudadas. Tal achado reforça quão influente as condições ambientais locais e globais são na conectividade e, por conseguinte, na diversidade genética das espécies.

Abstract

Understanding the main factors influencing connectivity among marine animal populations has been the subject of several studies for decades, with the mobility of individuals being recognized as one of the determining factors. Species with long larval development (e.g. planktotrophic) usually have better connected populations than species with direct or short larval development (e.g. lecithotrophic), due to their greater dispersal capacity. However, other bionomic features and the environment also have a major influence on connectivity. In addition, the perception of population connectivity can also be affected by the presence of cryptic species, since such species can have extreme morphological similarity. Thus, the present work evaluated how environmental heterogeneity influences the connectivity of species with different reproductive modes, through comparative landscape genomics techniques. For this purpose, we used the species *Perinereis ponteni* (Polychaeta, with sexual reproduction and lecithotrophic larvae), *Lineus sanguineus* (Heteronemertea, with asexual and sexual reproduction, and lecithotrophic larvae) *and Nemertopsis bivittata* (Hoplonemertea, with sexual reproduction, and presumably planktotrophic larvae). In order to guarantee the monophyly of all species, the populations studied were taxonomically assessed. In this sense, we sampled populations along the Brazilian coast. The collected animals were used in DNA extraction for SNPs prospection, using the Genotyping by Sequencing (GBS) technique. Furthermore, we followed the behavior and larval development of the hoplonemertean *Emplectonema viride* as a proxy for the development of *N. bivittata*, since both are from the Emplectonematidae family and live in extremely similar environments. However, the species *N. bivittata* is actually a complex of species, with three distinct evolutionary units on the Brazilian coast. These were described as, *Nemertopsis berthalutzae*, *Nemertopsis pamelaroeae* and *Nemertopsis caete*, the latter being present only on the coast of Alagoas. The observation of *E. viride* larval development indicated the presence of planktotrophic larvae, with long development, taking an average of 120 days for quite subtle metamorphosis (noticed mainly by changes in ciliation and behavior). The larvae feed primarily on barnacle nauplius and cyprids. Our palaeodistribution and demography analyzes indicate the presence of a refuge in the Southeast region during the Last Glacial Maximum, from where the populations expanded. As for landscape genomic analyses, we observed that species with shorter larval development tend to have more structured populations, but still with significant gene flow. However, similar environmental factors affect connectivity in all species, with temperature and precipitation being the most common. These factors are associated with the presence of SNPs candidates for natural selection in three of the four species studied. These findings reinforces how important local and global environmental conditions are for animal connectivity and, therefore, for the genetic diversity of species.