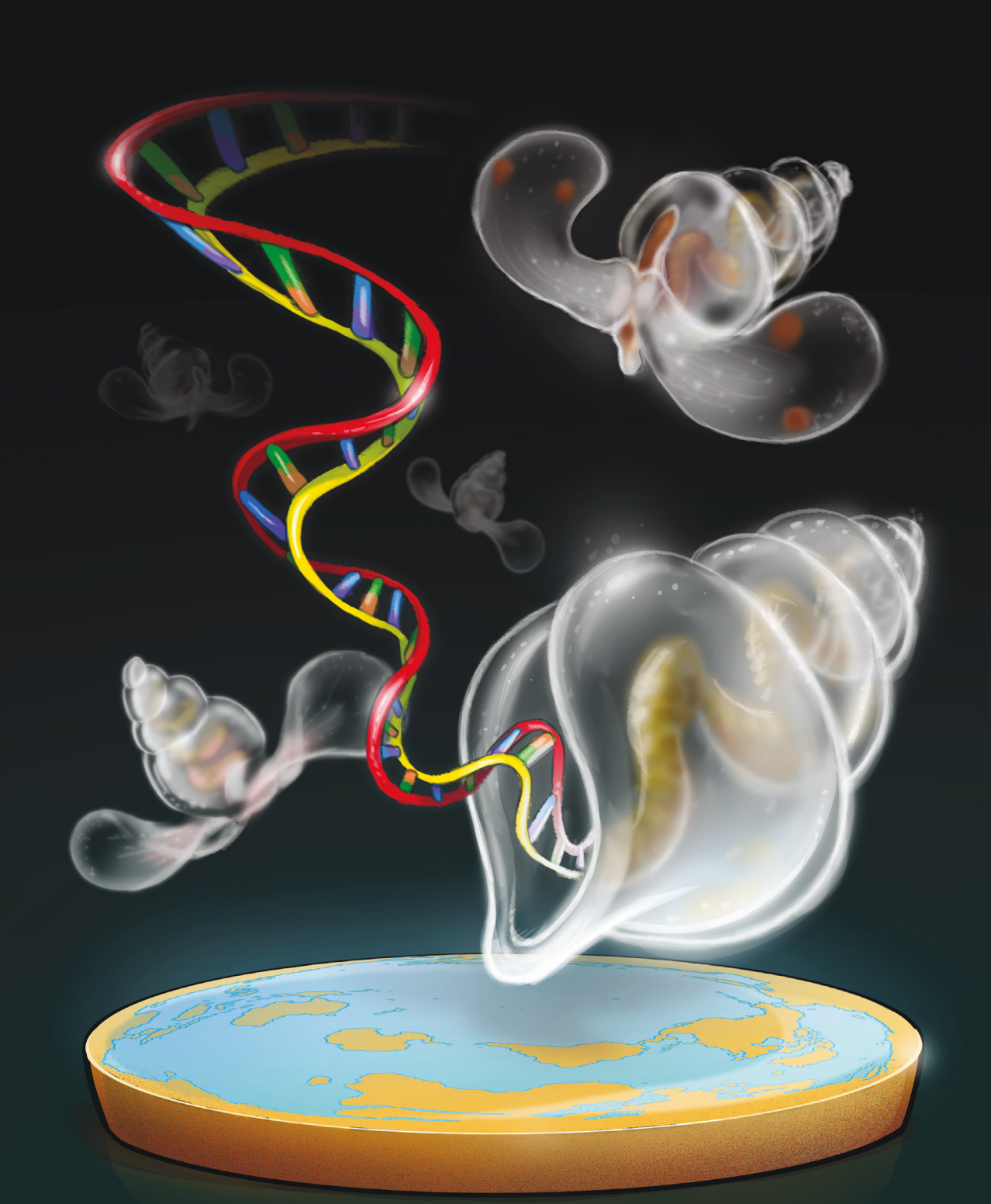
Genomic variability and population structure in shelled pteropods

L.Q. Choo



Genomic variability and population structure in shelled pteropods

Le Qin Choo

Genomic variability and population structure in shelled pteropods

Choo L.Q., 2022. Genomic variability and population structure in shelled pteropods

PhD thesis, University of Amsterdam, the Netherlands

This research was carried out at Naturalis Biodiversity Center, Leiden, the Netherlands, with research visits to Nord University, Bodø, Norway. The project was financially supported by the Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to Dr. Katja T.C.A. Peijnenburg.

ISBN: 978-94-93260-06-1

Cover design: Amalia Aikaterini Mailli, George Komiotis and Le Qin Choo
Thesis lay-out: Jan Bruin (www.bred.nl)

Genomic variability and population structure in shelled pteropods

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel

op donderdag 24 maart 2022, te 13.00 uur

door

Le Qin Choo

geboren te Singapore

Promotiecommissie

Promotores

prof. dr. J. Huisman	Universiteit van Amsterdam
prof. dr. G.G. Hoarau	Nord Universitet

Copromotor

dr. K.T.C.A. Peijnenburg	Universiteit van Amsterdam

Overige leden

prof. dr. K. Johannesson	Göteborgs universitet
prof. dr. M. Schilthuizen	Universiteit Leiden
dr. L.E. Becking	Wageningen University & Research
prof. dr. W. Renema	Universiteit van Amsterdam
prof. dr. A.T. Groot	Universiteit van Amsterdam
dr. S. Wilken	Universiteit van Amsterdam
prof. dr. C.P.D. Brussaard	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Table of contents

Author affiliations

Marine biodiversity, Naturalis Biodiversity Center, 2300 RA Leiden, the Netherlands

L.Q. Choo, K.T.C.A. Peijnenburg, P. Ramos-Silva, G. Spagliardi

Department of Freshwater and Marine Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, the Netherlands

L.Q. Choo, K.T.C.A. Peijnenburg, J. Huisman

Faculty of Biosciences and Aquaculture, Nord University, 8049 Bodø, Norway

T.M.P. Bal, M. Choquet, I. Smolina, M. Kopp, G. Hoarau

Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, USA

E. Goetze

Molecular Genetics Unit, Okinawa Institute of Science and Technology, Onna-son, 904-0495, Japan

F. Marlétaz

1

Introduction

Human activities are causing unprecedented changes in the ocean environment, which occupies 70% of the Earth's surface and is the largest habitat on Earth. Ocean temperatures are rising, while waters become less oxygenated and more acidic due to the emission of greenhouse gases. Plastic pollution accumulates in coastal regions and in ocean gyres, where they break down and enter the food chain. Humans are exploiting marine fish populations to the point of extinction, and are on the verge of exploiting the deep-sea for minerals on the ocean floor. We still know little about how these changes and impacts affect marine species and ecosystems. In addition, much is unknown about the ecology and evolution of many marine organisms. Will marine species be able to adapt to these changes and continue thriving? What will ecosystems look like in the future as species distributions and biodiversity patterns change? In this thesis, I focus on some key elements of these broad questions, zooming in on the genomic variability and population structure of shelled pteropods, a group of planktonic snails that inhabit the open ocean and are thought to be extremely sensitive to ocean acidification.

Since Darwin published The Origin of Species (Darwin, 1859), there has been much progress in our understanding of the factors and requirements influencing speciation. Speciation can be defined as the evolutionary process by which populations diverge and become reproductively isolated (i.e., are unable to interbreed with each other). We now know that the process of speciation can be mediated by a broad spectrum of mechanisms, depending on the extent of spatial separation and heterogeneity, and subsequent opportunities for gene flow between populations (Butlin et al., 2008; Mallet, 2008; Shaw and Mullen, 2014) (Figure 1). At one end of the spectrum is genetic divergence due to an extrinsic barrier, i.e., allopatry. In this case, populations are physically isolated from each other by an external barrier to dispersal, such as land bridges or mountain ranges. This leads to independent genetic drift and the accumulation of genetic differences, which results in the populations eventually being completely reproductively isolated. Allopatric speciation is often regarded as the most common form of speciation (Mayr, 1963), with most occurrences known from terrestrial examples where geographical barriers are easily identified. However, there are also examples of clear physical barriers in the marine environment, such as the Isthmus of Panama (Cowman and Bellwood, 2013). At the other end of the spectrum is sympatric speciation, where an initially interbreeding population splits into distinct species while still sharing the same geographic space at the same time, with the potential for interbreeding. Despite the theoretical support for this mode of speciation (Turelli et al., 2001; Via, 2001), there are relatively few well-supported examples of sympatric speciation for animals e.g., herbivorous insects (Bush, 1966), cichlid fish (Barluenga et al., 2006; Gavrilets et al., 2007) and seabirds (Friesen et al., 2007). For sympatric speciation to occur, genetic divergence has to proceed in the presence of homogenising gene flow, through mechanisms such as ecological divergence or assortative mating leading to reproductive isolation. Between the two ends of the spectrum, there are

speciation scenarios with partial genetic isolation and variable potentials for interbreeding and gene flow. Hence, it has been suggested that it is more appropriate to talk about a speciation continuum (Butlin et al., 2008). Adding to this complexity, it is important to note that different modes of speciation can occur more than once during the speciation process, which is not unidirectional or linear, but an accumulation of genetic divergences that can eventually lead to complete reproductive isolation (Stankowski and Ravinet, 2021). Studies of speciation in the ocean environment have been conducted on incipient or recently diverged species, mostly across their benthic or intertidal habitats (e.g., Bowen et al., 2013; Johannesson, 2009; Potkamp and Fransen, 2019; Stankowski et al., 2020) (Figure 1). Recently, with the increasing interest in the effects of climate change, researchers have also looked into the effect of environmental selection on the evolutionary divergence of marine species across their heterogeneous environment (Benestan et al., 2016; Davis et al., 2016; Morales et al., 2018; Xuereb et al., 2018b).

Speciation in the open ocean

Holoplanktonic species provide an interesting group of organisms to understand the process of marine speciation, because they drift with ocean currents their entire lives, in an environment that is characterised by the absence of obvious physical barriers and high connectivity across large distances. Previously, marine pelagic (i.e., open ocean) species were assumed to be panmictic, where all individuals of the population are able to interbreed with each other, due to their high potential for dispersal and gene flow (Norris, 2000; Palumbi, 1994; van der Spoel and Heyman, 1983). This was supported by the general finding of low genetic divergence in species across extensive geographical ranges (e.g., Apolônio Silva De Oliveira et al., 2017; Riginos et al., 2016; Selkoe et al., 2016). However, an increasing number of studies across pelagic taxa have uncovered genetic divergence in populations across a range of spatial scales (Andrews et al., 2014; Kulagin et al., 2021; Peijnenburg and Goetze, 2013; Pfaller et al., 2019; Postel et al., 2020; Truelove et al., 2017; Weersing and Toonen, 2009). This leads to an apparent paradox: speciation and population divergence should be unlikely in the open ocean since there are no obvious barriers to gene flow, but we still observe many populations that are in the process of diverging or have diverged. How species arise in the open ocean is thus an important open question (Filatov et al., 2021; Miglietta et al., 2011; Miya and Mishida, 1997; Peijnenburg et al., 2006; Peijnenburg and Goetze, 2013).

Analyses of genetic data have revealed that many open ocean species are more structured than would be expected based on their dispersal potential. Putative drivers of genetic differentiation in the open ocean have been identified based on case studies of pelagic organisms, although how these drivers interact depending on the ecology and evolutionary history of the organism is still relatively unknown. Population boundaries have been attributed to, for example, oceanographic barriers leading to limited connectivity and thus providing opportunities for allopatric speciation (Filatov et al., 2021), environmental transitions that limit the geograph-

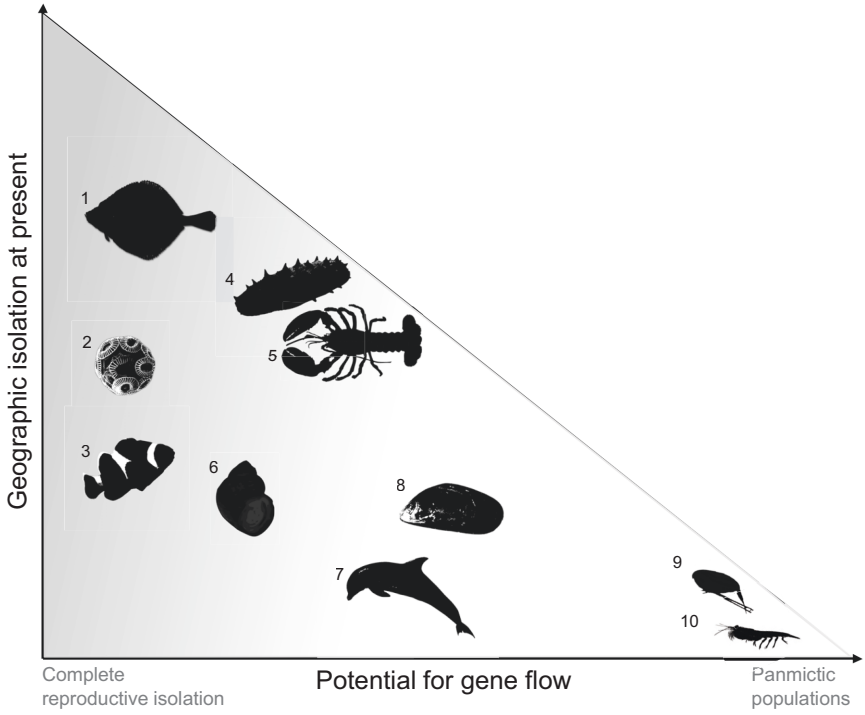


FIGURE 1 A schematic of the speciation continuum with some recent marine studies, showing genetic differentiation along the axes of current geographic isolation and the degree of gene flow. The degree of gene flow is indicated with the degree of grey shading from completely reproductively isolated species (grey) to panmictic populations (white). 1. European flounder in the Baltic Sea currently have an overlapping distribution, but are strongly reproductive isolated, likely due to ecological speciation resulting from different breeding behaviours (Momigliano et al., 2017). 2. Coccolithophore speciation in the open ocean most likely occurred through geographic isolation, and segregation of ecological niches, followed by present-day secondary contact (Bendif et al., 2019; Filatov et al., 2021). 3. Mutualistic interactions with host anemone species in clownfish can drive speciation through ecological selection, even in the absence of physical barriers (Litsios et al., 2012). 4. California sea cucumber exhibits local adaptation to environmental variables, which may play a role in spatially variable selection (Xuereb et al., 2018b). 5. Genetic differentiation in the American lobster is mediated by both thermal adaption and larval connectivity (Benestan et al., 2016). 6. Rough periwinkle (*Littorina saxatilis*) ecotypes (Johannesson, 2009; Johannesson et al., 2017) are partially reproductively isolated, with strong divergent selection, between sheltered coastal habitats facing crab predation pressure and exposed surfaces subject to wave action, keeping the ecotypes separate despite their physical proximity. 7. The Clymene dolphin (*Stenella clymene*) is a hybrid species of two putative parental species, *Stenella coeruleoalba* and *Stenella longisrostris*, all three of which occur in sympatry. 8. The *Mytilus* mussel species complex comprises multiple hybrid zones, with widespread local gene flow (Simon et al., 2021). 9. *Calanus finmarchicus* exhibits genetic homogeneity and gene flow across their broad geographic range in the northern North Atlantic, and hence are composed of panmictic populations (Choquet, 2017; Provan et al., 2009). 10. No population structuring was observed for Antarctic krill (*Euphausia superba*) across their range in the Southern Ocean (Deagle et al., 2015).

ical range of a species (Stanley et al., 2018), or a combination of these factors (Laso-Jadart et al., 2021). Complete geographic isolation is not necessary for the formation of these barriers, as there are documented examples of marine speciation in the presence of gene flow, as long as selective forces are strong enough to overcome the homogenising effect of gene flow (Bendif et al., 2019; Bierne et al., 2003; Potkamp and Fransen, 2019). Population boundaries can also arise depending on the interaction between the strength of environmental selection on dispersing plankton communities, and their rates of adaptation to non-optimal habitats (Ward et al., 2021). Given that marine plankton have enormous population sizes and relatively short generation times, they should be sensitive to even mildly beneficial selective forces (Peijnenburg and Goetze, 2013), but their actual extent of dispersal and selective pressure is still unknown.

Identifying and classifying species can be challenging in marine taxa, due to their large spatial ranges across an interconnected but relatively inaccessible habitat and their limited morphological differentiation. This can be seen from the various examples of circumglobally distributed species that were originally identified as a single species across their range due to morphological crypsis, only to be classified as separate species later due to genetic differences (Bongaerts et al., 2021; Hutchings and Kupriyanova, 2018; Knowlton, 1993, 2000). The problem is compounded when the organisms in question are small, relatively poorly studied, with scarce genetic resources and information about their morphology and ecology. Within the marine zooplankton, a large number of species have been reported as cryptic, however, for a large majority of examples there is incomplete evidence to ascertain that the species are actually reproductively isolated, and truly morphologically indistinguishable (van der Sprong, 2019). To definitively conclude that species exhibit cryptic diversity, a genome-wide perspective to identify instances of reproductive isolation, as well as in-depth analyses of morphological traits, are necessary. Overcoming these challenges will allow us to produce a more accurate representation of species ranges and their spatial genetic variability, which can then be used to gain insight into their potential to evolve in response to future change.

Effects of anthropogenic climate change on the ocean

The marine habitat has undergone multitudes of climatic changes in the past, with fluctuations in sea surface temperature, sea level and changes in ocean chemistry (e.g., Pelejero et al., 2010; Turney et al., 2020). The speed and extent of the current climate change is likely to be unmatched, however, with a rate of CO_2 release into the atmosphere that is unprecedented during past 66 million years (Zeebe et al., 2016). The effects of current ocean change are already visible in well-studied coastal ecosystems, for example, the images of bleached corals from the tropics are present in the public consciousness. Overall, ocean temperatures are expected to rise by 1-4 °C during the 21st century (Alexander et al., 2018; Gruber, 2011), while ocean waters become more acidic (Feely et al., 2004; Jiang et al., 2019) and less oxygenated (Keeling et al., 2010), with wide ranging and variable impacts on

different marine species and ecosystems (Doney et al., 2012, 2020; Fossheim et al., 2015; Gattuso et al., 2015; Hofmann et al., 2010; Kroeker et al., 2013).

Ocean acidification is one of the major changes that are happening to the ocean, and impacts a wide variety of marine taxa globally. The ocean absorbs about 30% of the anthropogenic carbon dioxide released annually into the atmosphere (Feely et al., 2009). With an increase in carbon dioxide emissions, the amount of carbon dioxide absorbed by the ocean will increase further. Carbon dioxide does not only dissolve but also reacts with seawater, releasing hydrogen ions (H^+), which causes seawater to become more acidic and carbonate ions to be less abundant. Carbonate ions are important building blocks for the calcium carbonate skeletons and shells of corals, molluscs, crabs and calcareous algae, hence, ocean acidification means that they will face difficulty in building and maintaining their shells and skeletons (Doney et al., 2012). In extreme cases of ocean acidification, these calcifying organisms face the prospect of dissolution (Bednaršek et al., 2012c; Hoegh-Guldberg et al., 2007; Riebesell et al., 2000). Hence, wide ranging habitats, from coral reefs to the open ocean, including polar and coastal seas are going to be affected, alongside other organisms within these ecosystems (Hofmann et al., 2010).

Marine species have shown various responses to climate change, such as range shifts, phenotypic plasticity and genetic adaptation. Range shifts have occurred across a wide range of species in response to changes in ocean temperature, with an overall poleward shift in planktonic species distributions (Beaugrand et al., 2009; Bedford et al., 2020; Benedetti et al., 2021; Garciá Molinos et al., 2016; Pinsky et al., 2020; Poloczanska et al., 2016). However, not all taxa exhibit range shifts that match the velocity of climate change, depending on whether they exhibit niche plasticity or conservatism, leading to changes in plankton communities over time (Chivers et al., 2017). Thus, range shifts can have wide-ranging impacts on marine biomass, species richness and ecosystem structure across ocean basins (Benedetti et al., 2021; Bryndum-Buchholz et al., 2019; Chaudhary et al., 2021).

Alternatively, species can also remain in their current geographic range and either possess the phenotypic plasticity to cope with changing conditions (Charmantier et al., 2008; Anderson et al., 2012; Barrett & Hendry, 2012), or adapt to the changing conditions (Benestan et al., 2016; Sunday et al., 2014). Phenotypic plasticity and adaptive evolution can also act together, with plasticity providing a buffer for the species to survive in the current environment, while selection acts on the gene pool so that the population becomes better suited to the changing environment (Chevin et al., 2010). A flipside of this is that plasticity can also slow down adaptive evolution, because it reduces the strength of selection by maintaining genetic variants that do not contribute to the adaptive phenotype (Sunday et al., 2014). There may also be ecological and genetic constraints associated with the limits to adaptation, and whether a population can adapt to environmental stress depends on many factors including population size, phenotypic variation, and demographic fluctuations (Bell, 2013; Chevin et al., 2013; Orr and Unckless, 2008; Tomasini and Peischl, 2020).

Marine zooplankton are key components of the marine food web, therefore, their ability to cope with future changes has an impact on other marine organisms, including phytoplankton, fish and marine mammals, as well as on biogeochemical fluxes. What do we know about the evolutionary potential of marine zooplankton? Theoretical arguments suggest that zooplankton should be able to adapt within short timeframes and in the presence of weak selection (Peijnenburg and Goetze, 2013). To accurately predict their capacity to cope with rapid environmental change, it is important to obtain information about their genetic variability and evolutionary potential.

Adaptive potential

Adaptive potential is defined as the ability of species and populations to respond to selection with phenotypic or genetic changes that improve their fitness (Eizaguirre and Baltazar-Soares, 2014). The adaptive potential of most marine species is difficult to measure directly because it is not possible to conduct breeding and long-term experiments with most non-model species. However, adaptive potential can also be indirectly inferred from sources of genomic information, including standing genetic variation, gene flow between populations, the strength of selection over time, and the demographic history of the species. Higher levels of standing genetic variation and the geographic context over which genetic variation is distributed is correlated with adaptive evolutionary responses (Miller et al., 2019; Ørsted et al., 2019). The demographic history can also provide a perspective into the levels of genetic variation across time, with a population bottleneck being associated with reduced genetic variation in present-day populations and enhanced susceptibility to the random, non-adaptive effects of genetic drift. For local adaptation to occur in a population, the effect of selection should exceed the homogenising effect of gene flow from adjacent populations, which can be achieved by a large effective population size, strong selection over time, or weak gene flow between populations (e.g., Attard et al., 2018). Local adaptation has been recorded for many marine species with a broad range of life histories, and understanding their spatial patterns of adaptation can give us insight into the potential impact of climate change on these species (Sanford and Kelly, 2011).

Pteropod biology and significance

This thesis focuses on shelled pteropods, a group of marine zooplankton that has been identified as extremely vulnerable to ocean acidification and which plays crucial roles in the global carbon cycle. Pteropods are a globally distributed group of gastropods, which are uniquely adapted to the open ocean habitat (Bé and Gilmer, 1977; van der Spoel and Dadon, 1999). The name ‘pteropod’ means wing-foot, referring to their foot, which has been modified into two wing-like flaps used for swimming in the water column. They are composed of two orders that are ecologically distinct, the Thecosomata (also known as sea butterflies), which graze on phytoplankton, bacteria and other microorganisms that are trapped in their mucus webs, and

the Gymnosomata (also known as sea angels), which are active predators, usually on the thecosomes (Figure 2). Within the Thecosomata, there are two groups, the Euthecosomata and Pseudothecosomata (Meisenheimer, 1905; Peijnenburg et al., 2020). The euthecosomes have either coiled or uncoiled calcium carbonate shells throughout their lives, while the pseudothecosome species have either coiled shells in both their larval and adult stages (the genus *Peracle*), or only a pseudoconch in their adult stages after discarding their larval shells (van der Spoel and Dadon, 1999). The gymnosomes lose their larval shells during metamorphosis into their naked adult forms (Lalli and Conover, 1976). Pteropods range in size between a few millimetres for the smallest thecosomes, to a few centimetres in the largest gymnosomes.

Despite their small size, shelled pteropods play pivotal ecological and biogeochemical roles. They are an important component of the pelagic food web globally, and are preyed upon by a wide range of taxa, including gymnosomes (Lalli and Gilmer, 1989), heteropods (Böer et al., 2005), amphipods (Bernard, 2006), cephalopods (Hanlon and Messenger, 1998), fishes (Hunt et al., 2008), and marine mammals (Lalli and Gilmer, 1989). In the polar and subpolar oceans, the gymnosome *Cliona limacina* has been observed to feed solely on the euthecosomatous pteropods *Limacina helicina* and *L. retroversa* (Hunt et al., 2008; Lalli and Gilmer, 1989). Pteropods also play a crucial role in the carbon flux across the global ocean (Bednaršek et al., 2012a), by sequestering carbon through grazing on phytoplankton (Hunt et al., 2008), contributing to the downward flux of carbon with their faecal pellets (Manno et al., 2010), mucus webs (Conley et al., 2018; Noji et al., 1997) and shells (Fabry et al., 2009; Steinberg and Landry, 2017; Tsurumi et al., 2005). Shelled pteropods also contribute up to 89% of total calcification in pelagic waters (Buitenhuis et al., 2019) and represent a major component of the calcium carbonate export from surface waters to the deep ocean (Sulpis et al., 2021).

The shells of euthecosomes are made of aragonite, a form of calcium carbonate that is about 50% more soluble than the other form, calcite (Mucci, 1983; Sun et al., 2015), and is thus more sensitive to dissolution under acidified conditions (Bednaršek et al., 2012b; Fabry et al., 2009; Manno et al., 2017; Orr et al., 2005). In addition, their shells are remarkably thin, ranging from a few µm in limaciniids to

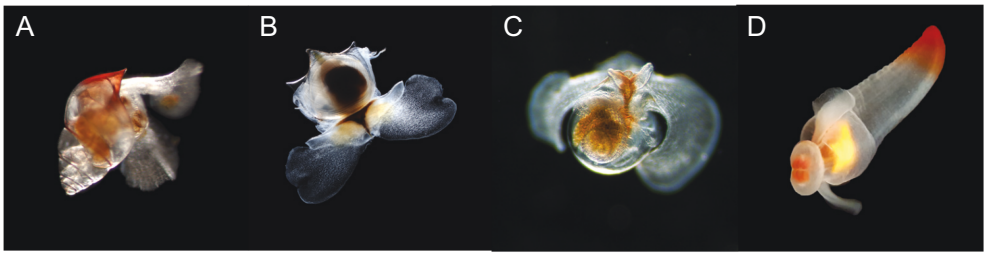


Figure 2 Example specimens for different groups of pteropods. (A) Coiled shelled euthecosome *Limacina bulimoides*, (B) Uncoiled shelled euthecosome *Cavolinia uncinata*, (C) Pseudothecosome *Peracle reticulata*, (D) Gymnosome *Clione limacina*. Images A-C by K.T.C.A. Peijnenburg & E. Goetze. Image D by L.Q. Choo and L. Mekkes.

100 µm in cavoliniids (Lalli and Gilmer, 1989). The shell morphology and size of thecosomes directly impacts their swimming and sinking behaviour (Karakas et al., 2020), which affects their efficiency in activities like feeding and predator avoidance through diel vertical migration. Diel vertical migration in zooplankton typically involves the movement of organisms to shallower waters of the ocean during the night, with a return to deeper waters during the day, as a means of minimising predation risk (Hays, 2003). This is not an insignificant migration, as pteropods ranging up to a few millimetres in size can routinely migrate between 50-200 m in depth every night (Karakas et al., 2020; Maas et al., 2012). Incubation experiments in which pteropods are exposed to future oceanic conditions and field observations in present-day acidified waters have both shown that the thin aragonitic shells of euthecosomatous pteropods are directly affected by acidifying conditions, by decreases in shell thickness and enhanced dissolution marks (Bednaršek et al., 2018; Busch et al., 2014; Mekkes et al., 2021a, 2021b; Niemi et al., 2021). Therefore, shelled pteropods are commonly regarded as important bioindicators for the effect of ocean acidification on marine calcifiers (Bednaršek et al., 2017b; Manno et al., 2017).

The presence of shelled pteropods in the fossil record also allows greater insight into the study of marine evolutionary dynamics. Pteropods are the only living metazoan plankton that is found consistently in the fossil record (e.g., Janssen, 2007, 2012; Wall-Palmer et al., 2014), and are reliable fossils for dating rock formations as their fragile nature means that they are rarely reworked from one sediment layer into another (Janssen and Peijnenburg, 2017). When combined with molecular phylogenetic inferences, the fossil record of pteropods can be used to calibrate the time of divergence between taxa (Burridge et al., 2017b; Corse et al., 2013; Peijnenburg et al., 2020).

Limacina genus

The most abundant shelled pteropod genus is *Limacina* (Bosc, 1817), a globally distributed genus spanning the polar to tropical oceans worldwide. The genus *Limacina* belongs to the superfamily Limacinoidea and is sister to the monotypic genus *Heliconoides*, which comprises the species *Heliconoides inflatus* (Figure 3).

Within the *Limacina* clade, five nominal species are currently accepted (Figure 4). Three warm-water species with (sub)tropical distributions:

- *Limacina bulimoides* (d'Orbigny, 1835)
- *Limacina trochiformis* (d'Orbigny, 1835)
- *Limacina lesueurii* (d'Orbigny, 1836)

Two cold-water species with bipolar or anti-tropical distributions:

- *Limacina helicina* (Phipps, 1774)
- *Limacina retroversa* (J. Fleming, 1823)

The cold-water species can be split further into their respective subspecies. *Limacina helicina* consists of four subspecies: *L. helicina ochotensis* Shkoldina, 1999, *L. helicina pacifica* Dall, 1871, *L. helicina rangii* (d'Orbigny, 1835) and *L. helicina*

helicina (Phipps, 1774). Two subspecies are recorded for *L. retroversa*: *L. retroversa australis* (Eydoux & Souleyet, 1840) and *L. retroversa retroversa* (Fleming, 1823).

Limacina species are characterised by their thin, aragonitic and sinistrally coiling shell and paired parapodia, which they use for swimming. The sea butterflies or 'papillon de mer' as coined by French fishermen in the eighteenth century (Lalli

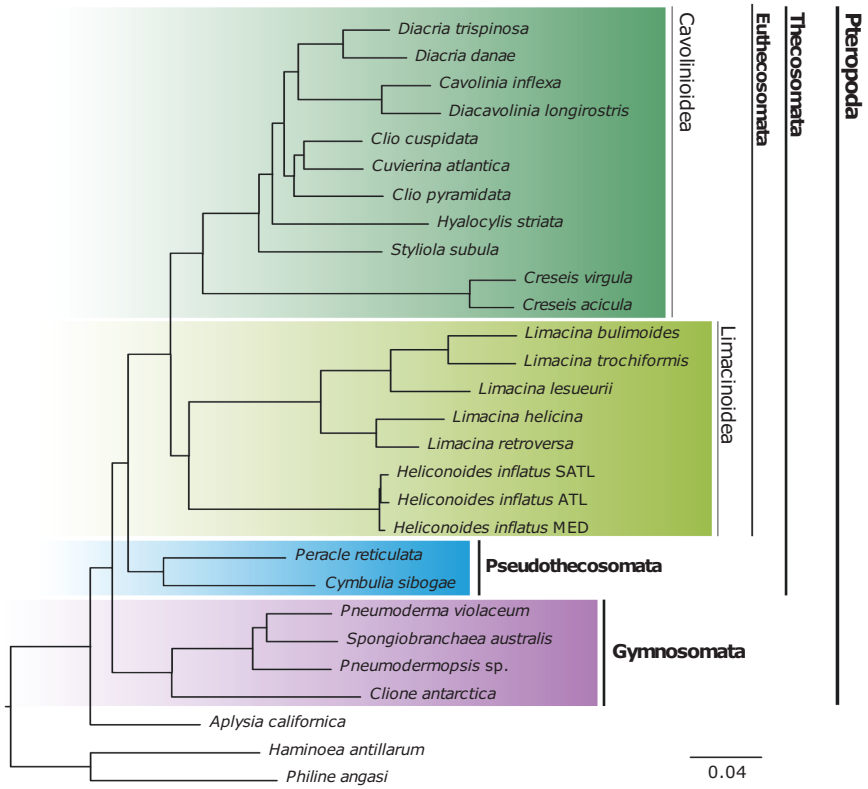


FIGURE 3 Phylogenomic tree resolving evolutionary relationships within pteropods. Maximum likelihood phylogeny of 25 pteropod taxa, plus 3 outgroups assuming an LG+Γ model. The dataset comprises 2,654 genes, concatenated as 834,394 amino acid positions with 35.8% missing data. Adapted from Figure 1 of Peijnenburg et al. (2020).

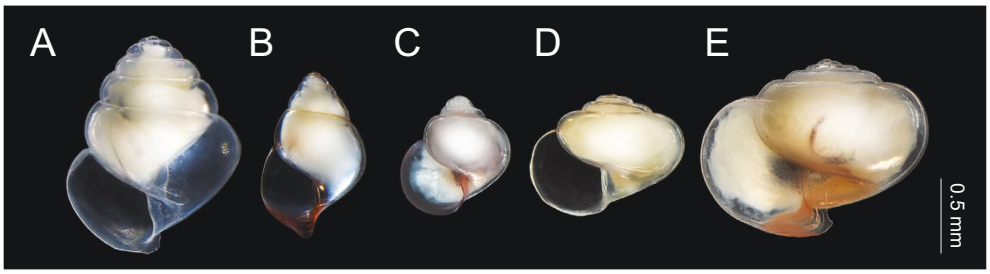


FIGURE 4 The five nominal *Limacina* species: (A) *L. retroversa*, (B) *L. bulimoides*, (C) *L. trochiformis*, (D) *L. lesueurii*, (E) *L. helicina*.

and Gilmer, 1989), are unique in terms of zooplankton locomotion. Pteropods have been shown to employ wing movements that are analogous to those used by many small, flying insects (Murphy et al., 2016), in contrast to other zooplankton that use their appendages as paddles (e.g., copepods and euphausiids). Their shell is negatively buoyant when wings are retracted, allowing them to sink rapidly for predator avoidance (Bergan et al., 2017; Gilmer and Harbison, 1986), and they attain neutral buoyancy when their wings are outstretched and parachute-like mucus webs are deployed (Gilmer and Harbison, 1986). They produce these mucus webs during feeding to capture planktonic organisms from the water column (Lalli and Gilmer, 1989). Their diet consists of phytoplankton, small protists and other particles trapped on their mucus webs, which they sort through with the use of ciliary pathways on their wings, footlobes and mantle lining prior to ingestion (Conley et al., 2018; Lalli and Gilmer, 1989). Species diversity in this genus is highest in the tropical and subtropical waters, while population density is highest in the (sub)polar waters (Burridge et al., 2017a; Lalli and Gilmer, 1989).

Biogeography of *Limacina* species

Limacina bulimoides is an abundant species found in circumglobal tropical and subtropical waters between 45°N-40°S, with the highest concentrations occurring in the central water masses, while typically being less abundant in tropical waters and boundary currents (Bé and Gilmer, 1977; Figure 5). In the Atlantic basin, *L. bulimo-*

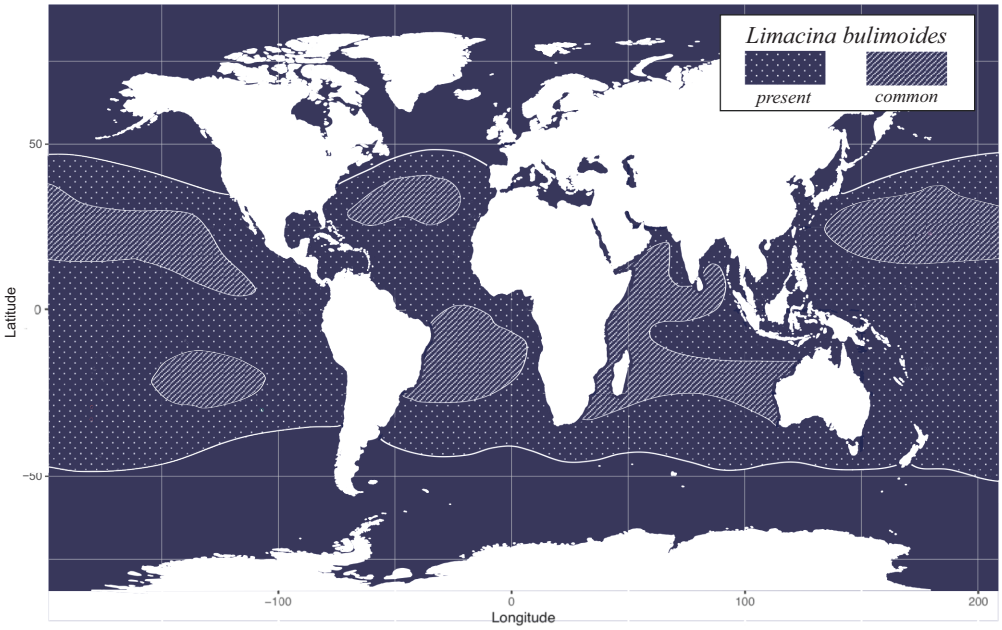


Figure 5 Global distribution of *Limacina bulimoides*, adapted from Bé and Gilmer (1977), which is based on live and sediment records. Dotted areas indicate the regions where the species has been recorded, while hatched areas indicate higher abundances of the species.

ides is found in the entire warm-water region from 45°N-40°S (Meisenheimer, 1905), and has been commonly recorded from the Sargasso Sea, the Gulf Stream (Tesch, 1946) and the Benguela current (van der Spoel and Dadon, 1999). A study along a meridional transect in the Atlantic Ocean by Burridge et al. (2017a) reported maximum peaks of abundance of ~100 individuals per 1000 m^3. A similar transect from three years later (in 2017) resulted in a maximum abundance of ~350 individuals per 1000 m^3 in the North Atlantic subtropical gyre (unpublished results). *Limacina bulimoides* is found between 40°N and 40°S in the Pacific, with a wide distribution across the Equatorial and Central water masses and in the Kuroshio Current. In the Indian Ocean, the highest concentrations of the species were recorded between 12°N and 10°S in the waters east of Somalia, and it was found to be scarce south of 30°S (Bé and Gilmer, 1977). *Limacina bulimoides* has a preferred depth range of 80-120 m (Bé and Gilmer, 1977), and has been documented to migrate vertically in the water column, to depths of 100-200 m during the day, and rising up to the upper 100 m of the water column at night (Wormuth, 1981).

Limacina lesueurii is a widespread and common subtropical species that is found in oligotrophic central water masses from 40°N to 40°S (Bé and Gilmer, 1977; Tesch, 1948; Figure 6). It is less abundant than its congener *L. bulimoides*, and appears to decrease dramatically in abundance near the equator (Bé and Gilmer, 1977; Meisenheimer, 1905) although these generalisations should also take into

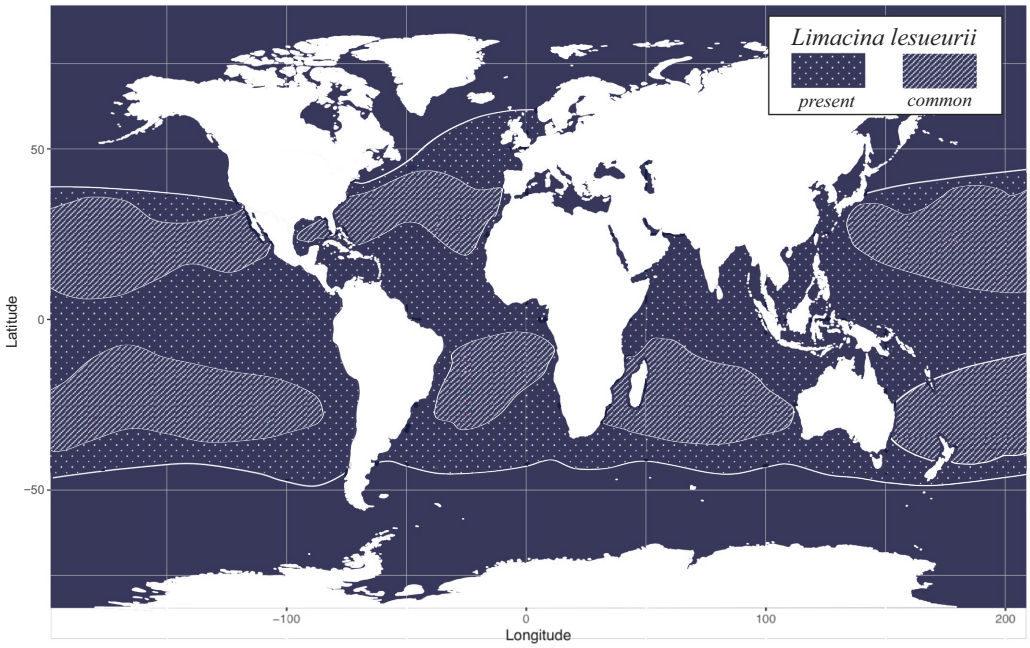


Figure 6 Global distribution of *Limacina lesueurii*, adapted from Bé and Gilmer (1977). Dotted areas indicate the regions where the species has been recorded, while hatched areas indicate higher abundances of the species.

account the patchy and possibly fluctuating abundance of this species. Burridge et al. (2017a) found a peak in abundance along an Atlantic meridional transect in the North Atlantic subtropical gyre, during autumn, of ~120 individuals per 1000 m^3. *Limacina lesueurii* is common in the western basin of the North Atlantic, except for the Gulf Stream (Bé and Gilmer, 1977). *Limacina lesueurii* is also abundant in the South Atlantic subtropical gyre (van der Spoel and Dadon, 1999). Within the Indian Ocean, this species is more common south of the equator, and especially in the Mozambique Channel (Bé and Gilmer, 1977), while in the Pacific Ocean it has been recorded from the upwelling region of the California current (McGowan, 1967, 1968), the Tasman Sea, as well as several sites off Pacific islands (Tesch, 1948). *Limacina lesueurii* demonstrates diel vertical migration (Wormuth, 1981). While it is considered an epipelagic species (usual depth range ~100 m), it has been recorded to be occasionally present at much greater depths up to 600 m in the Florida Current (Bé and Gilmer, 1977) and 1000 m in the Sargasso Sea (Wormuth, 1981).

Limacina trochiformis is a widespread warm-water species, and known to be most common in tropical upwelling regions, and coastal regions in the lower latitudes, but also reaches the higher latitudes due to transport by the boundary currents (Bé and Gilmer, 1977; Figure 7). In the Atlantic Ocean, *L. trochiformis* is most abundant in the Brazil Current, Florida Current and the Gulf Stream (Bé and Gilmer, 1977; van der Spoel and Dadon, 1999). Though Burridge et al. (2017a) reported

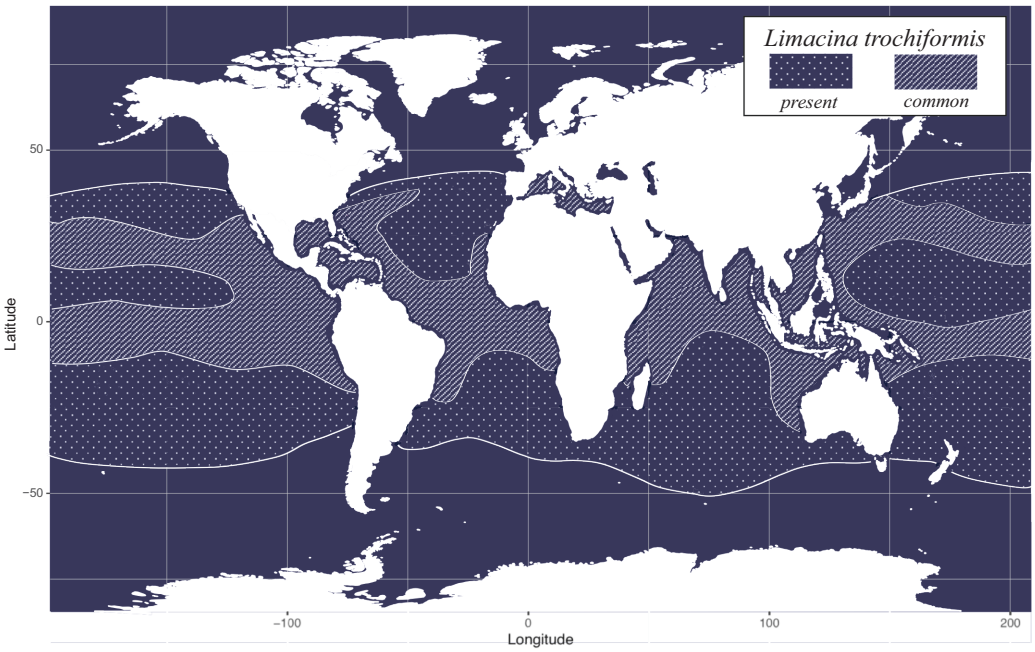


Figure 7 Global distribution of *Limacina trochiformis*, adapted from Bé and Gilmer (1977). Dotted areas indicate the regions where the species has been recorded, while hatched areas indicate higher abundances of the species.

low abundance of this species along the Atlantic meridional transect from 2014 (maximum abundance was 10 individuals per 1000 m^3), *L. trochiformis* was the most abundant *Limacina* species in the Brazil Current (Oliveira-Koblitz and Larrazábal, 2014). *Limacina trochiformis* is commonly found in the Indian Ocean, with peak occurrences off the coast of Somali between 0° and 10°N (Bé and Gilmer, 1977). In the Pacific Ocean, *L. trochiformis* has been recorded in the California Current, with fluctuating abundances through the year (McGowan, 1967, 1968). The vertical migration habits of *L. trochiformis* appear spatially variable, with vertical migration recorded in the Florida Current, but not off Cape Hatteras (Bé and Gilmer, 1977; Wormuth, 1981).

The nominal species *Limacina helicina* is present in the polar and sub-polar regions of the northern and southern hemispheres and shows a disjunct distribution pattern (Figure 8). While the southern subspecies, *L. helicina rangii* is currently recognised as a distinct species (WoRMS Editorial Board, 2021), the two species will be discussed together here as it is difficult to separate previous literature without additional genetic and taxonomic reviews. The *L. helicina* species complex comprises the northern hemisphere residents: *L. h. helicina*, *L. h. ochotensis*, *L. h. pacifica*, and the southern hemisphere *L. h. rangii* (McGowan, 1963; Shkoldina,

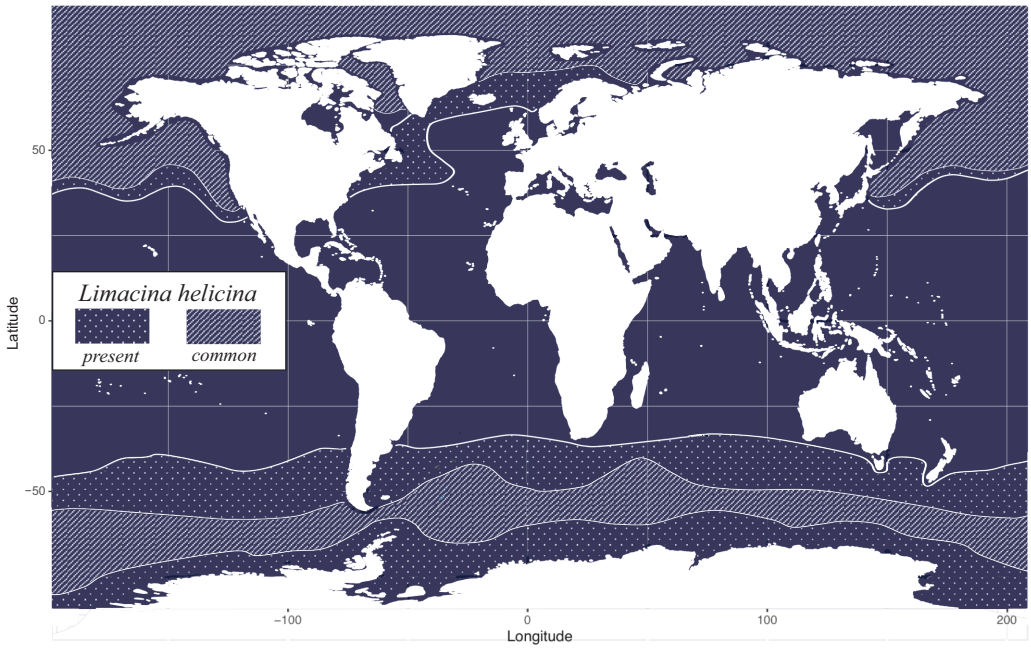


Figure 8 Global distribution of the *Limacina helicina* species complex, adapted from Bé and Gilmer (1977). Dotted areas indicate the regions where the species has been recorded, while hatched areas indicate higher abundances of the species. While the northern and southern populations are currently accepted as distinct species (WoRMS Editorial Board, 2021), they are represented together here as it is difficult to separate previous literature without additional genetic and taxonomic analysis.

1999; Tesch, 1948; van der Spoel, 1967; van der Spoel and Dadon, 1999). The range of *L. helicina* extends from the Arctic Ocean to sub-polar waters, up to 40°N in the Atlantic, and to temperate waters up to 25°N in the Pacific (Bé and Gilmer, 1977). It is currently unclear whether *L. rangii* or *L. h. rangii* should be the proper classification for this species, but substantial genetic differentiation has been observed between the Arctic *L. h. helicina* and Antarctic *L. h. rangii*, providing evidence that they are distinct species (Hunt et al., 2010; Peijnenburg et al., unpublished data). *Limacina helicina* can reach very high abundances averaging 165 individuals per m^3 south of the polar front in the Southern Ocean (Hunt et al., 2008). The species has been recorded as epipelagic, and was rarely found below depths of 300 m (Bé and Gilmer, 1977), with observations of diel vertical migration in *L. h. pacifica* (McGowan, 1963) but not in the southern hemisphere *L. h. rangii* (Foster, 1987).

Limacina retroversa is found in the sub-polar and transitional waters in the northern and southern hemispheres, but has not been recorded from the North Pacific (Bé and Gilmer, 1977; Figure 9). Within this species there are currently two accepted subspecies, *L. retroversa retroversa*, which inhabits the northern hemisphere and *L. r. australis*, which inhabits the southern hemisphere (van der Spoel, 1967). Both subspecies exhibit seasonal fluctuations in abundance and geographical range (Bé and Gilmer, 1977; Dadon, 1990; van der Spoel, 1967) but can reach very high abundances. For instance, in the Southern Ocean, the species was dominant north of the Polar Front, averaging 60 individuals per m^3 with a maximum of

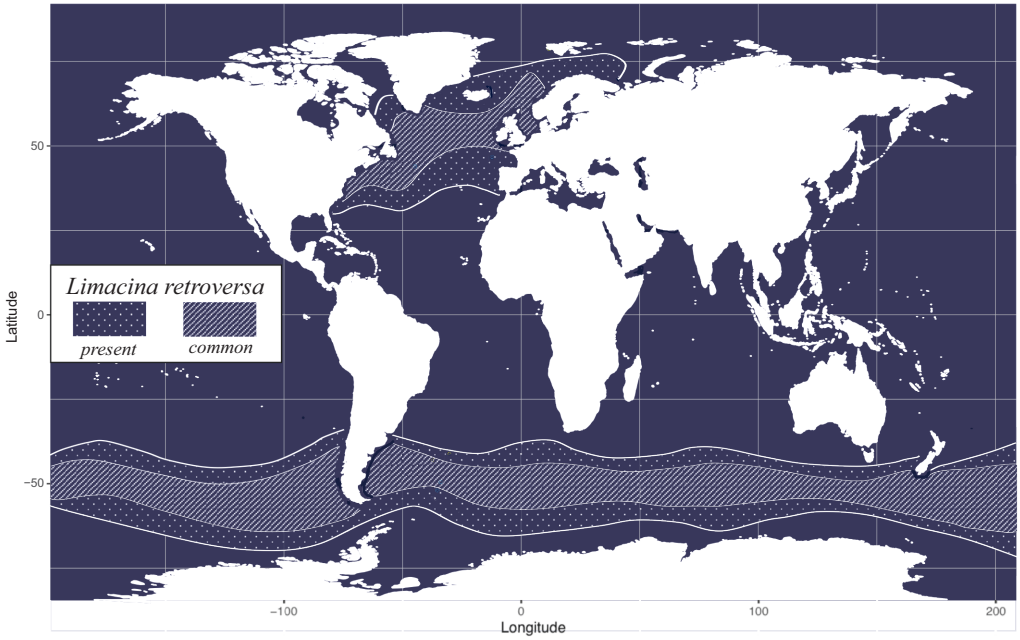


Figure 9 Global distribution of *Limacina retroversa*, adapted from Bé and Gilmer (1977). Dotted areas indicate the regions where the species has been recorded, while hatched areas indicate higher abundances of the species.

800 individuals per m^3 (Hunt et al., 2008). In the North Atlantic, *L. r. retroversa* can be found as far north as 81°N (Simal Busto, 2019), and in dense swarms in the Gulf Stream region between Newfoundland and the British Isles (Bé and Gilmer, 1977), while it is found in lower abundances in the Labrador Current, down to 20°N in the western Atlantic (Bé and Gilmer, 1977; Wormuth, 1985). This species has also been occasionally recorded in the Mediterranean basin (Bé and Gilmer, 1977) during glacial times. In the southern hemisphere, *L. r. australis* is found in a continuous band extending from the subtropical convergence at 38°S to a few degrees south of the Antarctic convergence and is the most abundant pelagic mollusc in the region (Bé and Gilmer, 1977; van der Spoel and Dadon, 1999). *Limacina retroversa* is an epipelagic species, and is most frequently found in the upper 150 m of the water column (Bé and Gilmer, 1977), although they seem to exhibit less pronounced diel vertical migration (Bernard and Froneman, 2009).

Knowledge gap and aims of this thesis

As explained above, shelled pteropods are species of interest for climate change research (Bednaršek et al., 2016, 2019; Keul et al., 2017; Manno et al., 2017). Many experimental studies have assessed their immediate response to future stressors, including warming, ocean acidification and reduction in food availability (Lischka et al., 2011; Maas et al., 2018; Thibodeau et al., 2020). However, these short-term experiments only give us information about their phenotypic plasticity, i.e., changes in behaviour, morphology or physiology that occur over their lifetime, but not of the adaptive potential of the population or species to persist into the future with environmental change. Experimental observation of adaptive changes through changes in allele frequency requires lab cultures across several generations. Such long-term experiments are feasible with microbes but impossible with shelled pteropods, because they are very challenging to maintain in the lab for a full life cycle (Howes et al., 2014) and because of their relatively long generation time of about a year. Therefore, information on shelled pteropods has to be obtained by studying natural populations, through observing the amount and distribution of genetic variation, population size, and degree of local adaptation. Based on these data, we can infer whether gene flow occurs between populations and at what spatial scales, and identify if local adaptation has occurred in the past, and is likely to occur in the future.

In this thesis, I focus on the shelled pteropod *Limacina bulimoides*, a highly abundant species with a widespread circumglobal warm-water distribution (Figure 5), which is well placed to answer questions about genetic structuring at different spatial scales. As with most other zooplankton species, it is unknown if this species is genetically homogeneous across its cosmopolitan distribution, or whether it may be composed of several cryptic species. Since most historical taxonomic efforts have categorised pteropods into species on the basis of shell shape or other identifiable morphological traits and not on the basis of genetic information (Burridge et al., 2015, 2019; Jennings et al., 2010), species boundaries may have well been overlooked. In fact, the genetic variability of *Limacina* species is largely unknown.

This thesis aims to characterise the genetic variability patterns in shelled pteropods of the *Limacina* genus, with the following objectives:

1. Design methods to assess genomic variability in *Limacina* species, including functional genetic variation
2. Assess the spatial distribution of genetic variation in *L. bulimoides* across ocean basins using DNA barcoding genes and genome-wide markers
3. Assess how genetic variation in *L. bulimoides* is related to morphological, ecological and environmental variables
4. Infer the potential of shelled pteropods to adapt to future changes based on present-day variability patterns

Approach

Shelled pteropods are mostly collected through oceanographic research cruises. A majority of pteropod samples used in this thesis are from the Atlantic Meridional Transect (AMT), a multidisciplinary programme consisting of annual cruises between the UK and destinations in the South Atlantic for the purpose of biological, chemical and physical oceanographic research. These transects spanned a wide range of biogeographic regions in the Atlantic Ocean, including the oligotrophic North and South Atlantic subtropical gyres and the more nutrient rich (i.e., mesotrophic) equatorial upwelling region. Samples from the 2012 and 2014 edition of the cruises formed the bulk of material studied for Chapter 2 (AMT22) and Chapter 5 (AMT24), respectively. In both these expeditions, bulk zooplankton samples were collected through nighttime oblique tows using a bongo net of 0.71 m diameter and mesh sizes of 200 and 333 µm or a RMT1 midwater trawl (333 µm) mesh size. Station metadata, including seawater temperature and chlorophyll concentration, were collected on site and calibrated by the British Oceanographic Data Centre (BODC). All samples used were preserved in 95% ethanol and stored at -20 °C to ensure optimal DNA preservation for subsequent genomic analyses.

Since *Limacina bulimoides* is a non-model organism, there are few genetic resources available. In pteropods, studies of genetic connectivity have been limited so far to molecular barcoding markers such as the mitochondrial cytochrome oxidase I (COI) and ribosomal 28S RNA genes (Burridge et al., 2015, 2019; Hunt et al., 2010; Shimizu et al., 2018, 2021). These barcoding markers have been used widely across zooplankton species (Bucklin et al., 2021b) and can provide easily accessible information on population structure without prior knowledge on the genome of the organism. However, the use of different markers can lead to discordant results, due to the haploid nature and uniparental inheritance of the mitochondrial genome, which can result in differing evolutionary histories between the mitochondrial and diploid nuclear genomes (Toews and Brelsford, 2012).

Genome-wide variation

With the recent advances and reduced costs of Next Generation Sequencing (NGS), it has become feasible to assess multiple regions across the genome (Allendorf et

al., 2010; Ekblom and Galindo, 2011). Sampling more sites across the genome provides greater resolution of the genetic population structure and barriers to dispersal by virtue of multiple observations, each providing a snapshot of the evolutionary history of the organism (Bucklin et al., 2018). For wild populations of non-model organisms, sampling large numbers of individuals across the entire metapopulation to get insight into their demography is not feasible. However, information into their demographic history can also be obtained with sampling many loci across the genome of a few individuals (Maisano Delser et al., 2016). Sequencing the whole genome of the organism is possible, especially for species with small genome sizes (<1GB); however, for a population scale sampling design of a zooplankton species with potentially large genome sizes (Bucklin et al., 2018), the costs are still prohibitive and amount of genetic information may quickly become intractable. Instead, it is more efficient to sample targeted loci across the genome, and increase the sequencing coverage to detect polymorphisms at those sites (Ekblom and Galindo, 2011).

Reduced representation sequencing (RRS) is an approach to generating genome-wide high-throughput sequencing data, by reducing the genomic data to be sequenced and NGS of the resulting genomic fragments. Currently, there are two main methods of RRS: restriction-site associated sequencing (RADseq) and targeted capture enrichment (discussed in the next paragraph). In RADseq, one or several restriction enzymes are used to cut DNA at predictable sites across the genome, and the regions flanking those restriction sites are selectively sequenced (Lowry et al., 2017). This method is attractive for non-model organisms, as no prior information about the genome is required (e.g., Blanco-Bercial and Bucklin, 2016; Deagle et al., 2015; Hirai, 2020). For RADseq, high amounts of DNA (up to 1 µg) are recommended (Etter et al., 2011), which may not be routinely feasible in small zooplankton with limited genetic material, although more recent modifications to RADseq protocols have shown that these methods can be implemented with much less DNA (50-100 ng) per individual (Andrews et al., 2016). Zooplankton typically have large, repetitive genomes, and one major drawback of RADseq is the inclusion of many anonymous fragments from across the genome, and a costly sequencing effort may be required to achieve sufficient coverage across all fragmented regions (Choquet et al., 2018b). Multicopy regions, which are present at a wide range of frequencies across the genome, are likely to be included and can result in a wide variation in coverage. Standard single nucleotide polymorphism (SNP) filtering protocols retain homologous SNPs with a minimum threshold coverage to ensure a confident SNP call, and exclude SNPs with excessive coverage, since they are likely to be highly repetitive regions. However, it can be difficult to distinguish between SNPs from homologous and repetitive regions if overall sequencing coverage of the RADseq dataset is low, e.g., because of a very large genome size (Deagle et al., 2015).

Target capture approach

An alternative method to RADseq is target capture, which is also known as targeted enrichment or hybridisation sequencing (Hyb-Seq). Target capture refers to the selective capture of genomic regions from a DNA sample before sequencing, through the use of pre-designed single-stranded DNA probes to bind with fragments of DNA from a mechanically sheared library, which are then recovered and amplified (Gnirke et al., 2009; Mamanova et al., 2010). Similar to RADseq, target capture has the advantage of being more cost- and time-effective as compared to whole genome sequencing, since less sequencing is needed, and the resulting datasets are less cumbersome to analyse. However, prior knowledge of the transcriptome or genome of the study organisms is required for the design of the target capture probes. Exon capture is a subset of target capture approaches, where probes are designed based on transcriptome assemblies of one or several species (e.g., Bi et al., 2012; Bragg et al., 2016; Portik et al., 2016). Exon capture circumvents the need for a genomic assembly, which may be complex and difficult to obtain in organisms with large and complex genomes. However, another approach, which mapped transcriptomic data to genomic data for producing genome-based target capture probes, resulted in better mapping quality than transcriptome-based capture probes, because of the inclusion of intron regions (Choquet et al., 2018b). While it requires additional effort to design probes based on prior knowledge of the genome and transcriptome, the benefits of target capture compared to RADseq include a high coverage because all loci across all individuals are the same, the ability to isolate specific sequences of interest, and a lower DNA input requirement (Chung et al., 2016; Jones and Good, 2016). Based on these benefits, genome-based target capture enrichment was chosen as the method for assessing genome-wide variability in *L. bulimoides*.

Morphological variation

While genetic and genomic markers provide a useful source of information about the evolutionary history of a species, additional information can also be gleaned from their morphological variation. Morphological variability provides a rich source of information about the selective pressures and physiological constraints acting on individuals, thereby yielding insights into patterns and processes leading to the distribution of biodiversity (Fišer et al., 2018). As such, integrative taxonomy approaches have been used in other shelled pteropod taxa to delineate species boundaries (Burridge et al., 2015, 2019; Shimizu et al., 2018, 2021). Since shell growth is accretionary, the shells of molluscs represent an ontogenic record of their life history, and provides insight into the adaptive constraints throughout their life (Vermeij, 2002). While shell shape variation can be captured by univariate measurements of shell dimensions (van der Spoel et al., 1993), geometric morphometrics is a powerful method for summarising multivariate variation in shell shape (Cruz et al., 2012; Roth and Mercer, 2000). In geometric morphometrics, shell shape is compared across individuals by the placing of landmarks on each shell image, followed by a Procrustes superimposition where landmarks are scaled and

rotated to remove differences in size and orientation, leaving only shape differences to be analysed by means of a principal components analysis (Rohlf and Slice, 1990). Since only a shell image is required as input for geometric morphometric analyses, a photograph of each pteropod individual in a standardised orientation was obtained before any destructive DNA extraction protocols.

Thesis outline

My thesis is structured as follows:

Chapter 2 assesses the spatial population structure of the pteropod species *Limacina bulimoides* in the Atlantic basin. For this purpose, we used two molecular markers, partial DNA sequences of the mitochondrial COI and nuclear 28S ribosomal barcoding genes, along with variation in shell shape using geometric morphometric analyses. The genetic and phenotypic variability was placed in an oceanographic context and linked to population abundances obtained from a related AMT transect to gain insight into the nature of oceanic dispersal barriers.

Chapter 3 focuses on the design of genome-wide target capture probes based on a draft genome and transcriptome of *L. bulimoides*. The target capture probes included putative biomineralisation genes, conserved pteropod orthologues, coding and non-coding regions, as well as commonly used DNA barcoding genes. We applied these genome-wide probes to the target species, *L. bulimoid*es, as well as four related shelled pteropod species: *L trochiformis*, *L. lesueurii*, *L. helicina*, and *Heliconoides inflatus* to assess the utility of these probes.

Chapter 4 elucidates the genome-wide population structure of *L. bulimoides* from the Atlantic, Indian and Pacific Ocean using the probes designed in Chapter 3. In addition, phenotypic variation including shell shape, shell colour and other morphological characters were examined. We identified the geographic distribution of different genetic lineages and gained insight into historical processes that could have influenced their present-day population structure and abundance.

Chapter 5 builds upon Chapter 4 and investigates the finer-scale population structure of *L. bulimoides* in the Atlantic basin. We compared the inferences of population structure in the Atlantic Ocean derived from genome-wide markers versus barcoding genes, to gain deeper insight into the drivers of spatial and temporal population structure in this holoplanktonic species. In particular, we identify three genetically distinct populations of *L. bulimoides*, inhabiting the North, Equatorial, and South Atlantic, respectively.

Inspired by the COVID pandemic, Chapter 6 explored and recommends the use of alcohol-based hand sanitiser as a medium for the morphometric and genetic analyses of shelled pteropods, and potentially other small (planktonic) organisms. Hand sanitiser increases positioning accuracy and efficiency in stacking microscope photography. This facilitates the inclusion of such photographs for voucher specimens that are challenging to photograph in a standardised orientation due to their small size, and enables their inclusion in reference databases.

Chapter 7 integrates the findings from Chapters 2 to 6 to address the concept of genetic structuring in the open ocean, with the pteropod *L. bulimoides* as a focal example. I show that, in comparison to traditional barcoding genes, genomic data provide more detailed insight into the drivers of genetic structure between populations and species. In the absence of clear hydrographic barriers, *L. bulimoides* exhibits geographically separated, evolutionarily independent lineages and populations between and within ocean basins.

2

Oceanic dispersal barriers in a holoplanktonic gastropod

L.Q. Choo, T.M.P. Bal, E. Goetze, K.T.C.A. Peijnenburg

Abstract

Pteropods, a group of holoplanktonic gastropods, are regarded as bioindicators of the effects of ocean acidification on open ocean ecosystems, because their thin aragonitic shells are susceptible to dissolution. While there have been recent efforts to address their capacity for physiological acclimation, it is also important to gain predictive understanding of their ability to adapt to future ocean conditions. However, little is known about the levels of genetic variation and large scale population structuring of pteropods, key characteristics enabling local adaptation. We examined the spatial distribution of genetic diversity in the mitochondrial cytochrome *c* oxidase I (COI) and nuclear 28S gene fragments, as well as shell shape variation, across a latitudinal transect in the Atlantic Ocean (35°N-36°S) for the pteropod *Limacina bulimoides*. We observed high levels of genetic variability (COI $\pi = 0.034$, 28S $\pi = 0.0021$) and strong spatial structuring (COI $\Phi_{ST} = 0.230$, 28S $\Phi_{ST} = 0.255$) across this transect. Based on the congruence of mitochondrial and nuclear differentiation, as well as differences in shell shape, we identified a primary dispersal barrier in the southern Atlantic subtropical gyre (15-18°S). This barrier is maintained despite the presence of expatriates, a gyral current system, and in the absence of any distinct oceanographic gradients in this region, suggesting that reproductive isolation between these populations must be strong. A secondary dispersal barrier supported only by 28S pairwise Φ_{ST} comparisons was identified in the equatorial upwelling region (between 15°N-4°S), which is concordant with barriers observed in other zooplankton species. Both oceanic dispersal barriers were congruent with regions of low abundance reported for a similar basin-scale transect that was sampled two years later. Our finding supports the hypothesis that low abundance indicates areas of suboptimal habitat that result in barriers to gene flow in widely-distributed zooplankton species. Such species may in fact consist of several populations or (sub)species that are adapted to local environmental conditions, limiting their potential for adaptive responses to ocean changes. Future analyses of genome-wide diversity in pteropods could provide further insight into the strength, formation and maintenance of oceanic dispersal barriers.

Keywords

shelled pteropods, zooplankton, geometric morphometrics, phylogeography, Atlantic Ocean

This chapter is published as

Choo, L.Q., Bal, T.M.P., Goetze, E., Peijnenburg, K.T.C.A., 2021. Oceanic dispersal barriers in a holoplanktonic gastropod. J. Evol. Biol. 34, 224–240.

Introduction

Anthropogenic carbon emissions are an important cause of perturbations in marine ecosystems, including an overall increase in seawater temperature, ocean deoxygenation and acidification (Gattuso et al., 2015; Gruber et al., 2019). These perturbations are expected to have wide-ranging impacts on marine organisms and may drive species range shifts and ecosystem regime shifts due to climate-driven invasions and extinctions (Doney et al., 2012; Kroeker et al., 2013; Poloczanska et al., 2016). It is uncertain whether marine organisms have the short-term plasticity and/or long-term evolutionary potential to adapt, given the ongoing and expected rates of environmental change (Donelson et al., 2019; Miller et al., 2018). To gain insight into the evolutionary potential of marine taxa, it is necessary to resolve their underlying genetic variation, population structure and connectivity (Bell, 2013; Harvey et al., 2014; Munday et al., 2013; Poloczanska et al., 2016; Sunday et al., 2014).

Locally adapted populations are increasingly reported in marine systems, showing that the apparent connectivity of the marine environment is often not fully realised by organisms, and populations may be more spatially constrained than once thought (Barth et al., 2017; Nielsen et al., 2009; Sanford and Kelly, 2011). Several case-studies have illuminated the processes involved in local adaptation of marine molluscs, e.g., habitat preference and partial spawning asynchrony in the mussels *Mytilus edulis* and *M. galloprovincialis* (Bierne et al., 2003), ecological selection across microhabitats in the ecotypes of the rocky shore gastropod *Littorina saxatilis* (Butlin et al., 2014; Johannesson et al., 2017; Westram et al., 2018), or variation in thermal stress tolerance in the intertidal gastropod *Chlorostoma funebralis* (Gleason and Burton, 2016). Much of the existing literature has focused on dispersal barriers in marine benthic taxa, some of which have pelagic larval stages. However, little is known about evolutionary pressures acting in the pelagic environment. One reason is that it is difficult to differentiate between selection pressures acting on the larval pelagic and adult benthic life stages on the same genome (Marshall and Morgan, 2011). Thus, it could be useful to study wholly pelagic organisms, such as holozooplankton, to better understand the selective forces acting upon organisms in the open ocean.

Population structure in marine holozooplankton appears to be more influenced by their ability to establish and maintain viable populations outside of their core range, rather than dispersal limitations (De Vargas, Norris, Zaninetti, Gibb, & Pawlowski, 1999; Norton and Goetze, 2013; Peijnenburg, Breeuwer, Pierrot-Bults, & Menken, 2004). Given the large population sizes, extensive distribution patterns, high standing genetic diversity, and short generation times of marine zooplankton in general, adaptive responses to even weak selection may be expected as the loss of variation due to genetic drift should be negligible (Peijnenburg and Goetze, 2013). While some plankton communities show rapid range shifts and phenological changes in response to ocean warming (Beaugrand, Goberville, Luczak, & Kirby,

2014; Beaugrand, Luczak, & Edwards, 2009; Hays, Richardson, & Robinson, 2005), we only have limited understanding of the potential for evolutionary responses of marine zooplankton to future conditions (Dam, 2013; Peijnenburg and Goetze, 2013).

Pteropods are a group of holoplanktonic gastropods that play important roles in planktonic food webs and ocean biogeochemistry, due to their high abundance and production of calcium carbonate shells (Bé and Gilmer, 1977; Bednaršek et al., 2012a; Buitenhuis et al., 2019; Hunt et al., 2008). Shelled pteropods are especially vulnerable to global change, because of their thin aragonitic shells that are prone to dissolution (Bednaršek et al., 2012c; Lischka et al., 2011; Manno et al., 2017), and they have been suggested as bioindicators of the effects of ocean acidification (Bednaršek et al., 2017). Several prior genetic studies on shelled pteropods have focused on resolving species boundaries via an integrative approach, combining genetic analyses and morphometric measurements of shells, to obtain a better understanding of species distribution patterns (Burridge et al., 2015, 2019; Shimizu et al., 2018). Other recent research efforts have addressed the response of shelled pteropods to acidified conditions to evaluate their acclimation to rapid increases in ocean acidity (Bednaršek et al., 2017a; Bogan et al., 2020; Maas et al., 2018; Moya et al., 2016). Little is known, however, about the spatial distribution of natural genetic and phenotypic variability in pteropod species with widespread distributions.

Planktonic ecosystems in the Atlantic Ocean have been relatively well-sampled owing to spatially extensive (~13,500 km) annual transects of the Atlantic Meridional Transect programme (AMT, https://www.amt-uk.org), and are thus ideal for studying population structure and dispersal barriers in marine zooplankton. The AMT cruises traverse both the northern and southern subtropical gyres, which are separated by the equatorial upwelling region. The subtropical gyres are oligotrophic systems characterised by clear ocean waters with very low primary productivity in the surface layer, high sea surface temperature, a deep thermocline, and the presence of a deep chlorophyll maximum. Conversely, the mesotrophic equatorial province is characterised by upwelling of nutrient-rich deep waters that stimulate primary production. Several species of copepods show spatial genetic structuring congruent with these Atlantic oceanic provinces. For example, *Pleuromamma abdominalis* has clades that are endemic to the equatorial province (Hirai et al., 2015). Other mesopelagic copepods, such as *Haloptilus longicornis* (Andrews et al., 2014; Norton and Goetze, 2013) and *Pleuromamma xiphias* (Goetze et al., 2017) show strong genetic breaks between the northern and southern subtropical gyre populations. These dispersal barriers coincide with regions of low abundance, which may represent areas of suboptimal habitat, supporting an ecological basis for these barriers (Goetze et al., 2015, 2017). However, it is unclear whether other zooplankton groups show similar patterns of basin-scale genetic structuring. For holoplanktonic gastropods, DNA barcoding has revealed intra-specific genetic differentiation between and within ocean basins,

pointing towards undescribed diversity (Burridge et al., 2017b; Jennings et al., 2010; Wall-Palmer et al., 2018). Using DNA barcoding, two closely related species of *Protatlanta* (a genus of Atlantidae, commonly known as shelled heteropods) were identified in the Atlantic Ocean: *P. souleyeti*, which is abundant in the subtropical gyres, and *P. sculpta*, which is most abundant in the equatorial upwelling region (Wall-Palmer et al., 2016a). There is also some evidence that populations of the straight-shelled pteropod *Cuvierina atlantica* show genetic discontinuity across the equatorial upwelling region in the Atlantic Ocean (Burridge et al., 2015).

In this study, we aim to identify barriers to dispersal and spatial population structure in the coiled-shelled pteropod *Limacina bulimoides* (d'Orbigny, 1835) across a latitudinal transect in the Atlantic Ocean. This species has a circumglobal warm-water distribution from ~45°N to ~40°S, a preferred depth range of 80 to 120 meters, and it performs diel vertical migration with higher abundance in surface waters at night (Bé and Gilmer, 1977). In the Atlantic basin, the species is common across several ocean provinces with peaks in abundance in the oligotrophic gyres (Bé and Gilmer, 1977; Burridge et al., 2017a). By assessing the distribution of genetic and phenotypic variability in *L. bulimoides* across a latitudinal Atlantic transect, we aim to test the hypothesis that areas of low abundance mark regions of suboptimal habitat that represent barriers to gene flow. To do this, we sequenced fragments of the mitochondrial cytochrome *c* oxidase I (COI) and the nuclear 28S genes and made geometric morphometric measurements of shell shape. Our objectives were to (1) identify the location of dispersal barriers based on two genetic markers, (2) test for congruence of genetic barriers with differences in shell shape, and (3) assess whether the spatial structure was better explained by shifts in abundance or oceanographic parameters. By identifying dispersal barriers and the possible drivers of population structure in holoplanktonic gastropods, we can better predict their capacity to respond to a rapidly acidifying ocean.

Materials and Methods

Species ecology and biology

Limacina bulimoides is a sinistrally-coiled holoplanktonic gastropod, well-adapted to life in the open ocean. It has a small (maximum shell length = 2 mm; van der Spoel et al., 1997), thin, aragonitic shell and two parapodia, which it uses to swim through the water column in analogous fashion to how insects fly through the air - a remarkable example of convergent evolution (Murphy et al., 2016). For feeding, an external mucous web traps particulate matter, including phytoplankton and small protists, which is brought to the mouth by ciliary movement (Lalli and Gilmer, 1989). The radula, a typical molluscan structure used for feeding, has been described from a congener, *L. helicina*, and was reduced, with 10 rows of teeth (Gilmer and Harbison, 1986; Lalli and Gilmer, 1989; Meisenheimer, 1905; Ritcher, 1979). However, in a recent detailed morphological study of the same species, only six rows of teeth were found, which led the authors to suggest that the number of rows could be variable

due to continuous regrowth (Laibl et al., 2019). *Limacina bulimoides* lay free-floating egg strings (see Figure S1/Supplementary Video 1), from which free-swimming veliger larvae hatch (Lalli and Wells, 1978). *Limacina* spp. are protandrous hermaphrodites, hence, the larvae sexually mature into males before transitioning into females (Lalli and Gilmer, 1989). The generation time of *L. bulimoides* is estimated to be less than a year, with larvae metamorphosing into juveniles after 2 months, juveniles reaching sexual maturity as males after a subsequent 5 months, and reaching their maximum size as females 2 months later (Wells, 1976). Reciprocal copulation has been observed between males (Morton, 1954), or between individuals that are in transition between male and female (Lalli and Gilmer, 1989). They possess a penis for internal fertilization (Lalli and Wells, 1978), and it is unknown whether self-fertilisation is possible.

Sampling

Bulk plankton samples were collected on AMT cruise 22 (AMT22) in October and November 2012 (Table 1, Figure 1). The samples were collected at night by oblique tows of a bongo net (200 and 333 µm mesh sizes) or RMT1 midwater trawl (333 µm) from a median depth of 323 m to the sea surface (depth range: 132 - 402 m). Samples were preserved in 95% ethanol, stored at -20 °C, and subsequently sorted in the laboratory. Seawater temperature and chlorophyll *a* concentration in the upper 300 m of the water column were obtained using a Sea-Bird Electronics 3P Temperature Sensor and Chelsea MKIII Aquatracka Fluorometer, with data calibrated and archived by the British Oceanographic Data Centre (BODC, https://www.bodc.ac.uk). The 14 samples included in this study were collected between 35°N and 36°S in the following Longhurst biogeochemical provinces (Longhurst, 2007; Reygondeau et al., 2013): North Atlantic Subtropical gyre (NAST), North Atlantic Tropical gyre (NATR), Western Tropical Atlantic (WTRA), South Atlantic gyre (SATL), South Subtropical Convergence (SSTL) (Table 1). Additionally, one sample site of *L. bulimoides* from the Pacific was included as an outgroup to provide a broader perspective on the diversity found for the Atlantic individuals (Table 1, Figure 1). For population-level analyses, the Longhurst province assignments were simplified to four major ocean provinces (North Gyre, Equatorial, South Gyre and Convergence) (Table 1), which also take into account previously reported transitions in planktonic community composition along the AMT transects (Burridge, Goetze, et al., 2017; Burridge, Tump, Vonk, Goetze, & Peijnenburg, 2017; Goetze et al., 2017).

Abundance of *L. bulimoides* was obtained from AMT24 (September-October 2014) as quantitative samples from AMT22 were not available. For this equivalent AMT transect, bulk plankton samples were collected using a bongo net (200 µm mesh size) with a General Oceanics flowmeter (2030RC) mounted in the mouth to quantify the volume of seawater filtered (Burridge et al., 2017). Based on the peaks in abundance of *L. bulimoides* across the AMT24 transect and comparison of oceanographic measurements from both the AMT22 and AMT24 transects, the sampled stations were split into three 'population groups' (North, Equatorial, South) (Table 1).

Table 1 Sampling overview for *Limacina bulimoides* from the Atlantic AMT22 (14 stations) and Pacific KH11-10 (1 station) cruises, including the location, date collected, Longhurst province assignment of the stations, number of individuals sampled for mitochondrial cytochrome *c* oxidase I (COI), nuclear 28S (28S) genes, and morphometric data. The Longhurst codes mentioned in the table are as follows: NAST = North Atlantic Subtropical gyre, NATR = North Atlantic Tropical gyre, WTRA = Western Tropical Atlantic, SATL = South Atlantic gyre, SSTC = South Subtropical Convergence, NPTG = North Pacific Tropical gyre. Sites with two Longhurst province assignments indicate transitional zones. A division of sampling sites into groups representing four major ocean provinces and three groups based on abundance along a similar transect are indicated in the last two columns, respectively.

Cruise and station	Latitude	Longitude	Date	Longhurst code	COI	28S	Morpho-metrics	Ocean provinces	Population groups
AMT22_13	34°22'N	27°38'W	18/10/2012	NAST/NATR	28	28	10	North Gyre	North
AMT22_19	27°36'N	36°22'W	21/10/2012	NATR	31	26	17	North Gyre	North
AMT22_23	23°9'N	40°38'W	23/10/2012	NATR	28	24	2	North Gyre	North
AMT22_25	20°24'N	38°37'W	24/10/2012	NATR	19	19	21	North Gyre	North
AMT22_29A	15°18'N	34°40'W	26/10/2012	NATR/WTRA	21	19	20	Equatorial	North
AMT22_43A	4°19'S	25°1'W	1/11/2012	WTRA	25	20	3	Equatorial	Equatorial
AMT22_45	8°5'S	25°2'W	3/11/2012	WTRA/SATL	41	41	6	Equatorial	Equatorial
AMT22_49	15°18'S	25°4'W	5/11/2012	SATL	27	31	-	South Gyre	Equatorial
AMT22_51	18°30'S	25°6'W	6/11/2012	SATL	16	16	1	South Gyre	South
AMT22_53	20°6'S	24°31'W	6/11/2012	SATL	25	25	17	South Gyre	South
AMT22_55	22°57'S	25°0'W	8/11/2012	SATL	38	38	6	South Gyre	South
AMT22_60	30°10'S	27°54'W	12/11/2012	SATL/SSTC	24	28	8	Convergence	South
AMT22_62	34°7'S	33°30'W	14/11/2012	SATL/SSTC	8	7	-	Convergence	South
AMT22_64A	35°52'S	36°0'W	14/11/2012	SATL/SSTC	3	2	-	Convergence	South
KH1110_08	22°47'N	158°06'W	19/12/2011	NPTG	22	20	25	–	–
Total					356	344	136		

Molecular analyses

To resolve spatial genetic structuring of *L. bulimoides*, 356 individuals were sequenced for a 565 base pair (bp) fragment of the mitochondrial COI gene, and 344 individuals for a 938 bp fragment of the nuclear 28S gene (Table 1). In total, 332 *L. bulimoides* specimens were sequenced for both COI and 28S (Table S1). The average number of individuals sampled per station for COI and 28S was 24 and 23, respectively. Individuals from stations 62 and 64 were combined because of their geographic proximity and low numbers of specimens. DNA was extracted with either the NucleoMag 96 Tissue Kit (Macherey-Nagel GmbH & Co. KG) or the DNeasy 96 Blood & Tissue Kit (Qiagen). DNA extracts were amplified in polymerase chain reaction (PCR) using the primers jgLCO1490 (5'-TITCIACIAAYCAYAARGAYATTGG-3') and jgHCO2198 (5'-TAIACYTCIGGRTGICCRAARAAYCA-3') for COI (Geller et al., 2013), and for 28S using C1-F (5'-ACCCGCTGAATTTAAGCAT-3') (Dayrat et al., 2001) and D3-R (5'-GACGATTCGATTTGCACGTCA-3') (Vonnemann et al., 2005). PCR was carried out with a reaction mix consisting of 1 µl template DNA, 17.8 µl milli-

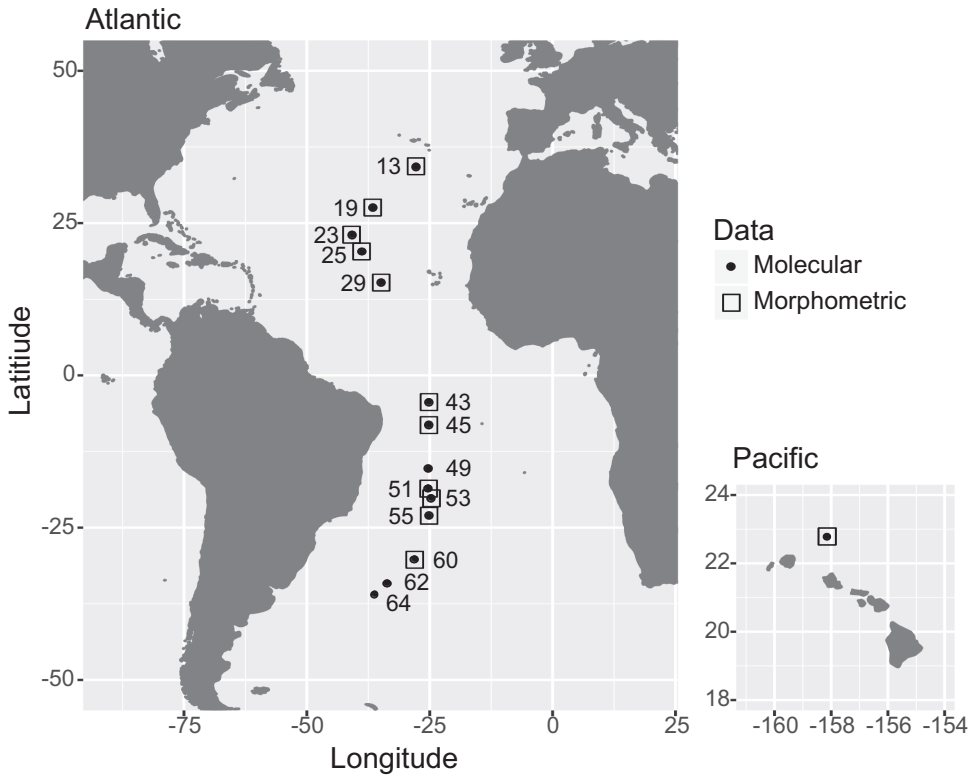


Figure 1 Sampling locations of *Limacina bulimoides* along the basin-scale Atlantic Meridional Transect (AMT22, left), as well as in the Pacific Ocean off Hawai'i (right). Numbers indicate the station number for each site, and symbols indicate the type of data that was obtained (see legend).

Q H_2O (Ultrapure), 2.5 µl PCR buffer CL (10X) (Qiagen), 0.5 µl $MgCL_2$ (25 mM) (Qiagen), 0.5 µl BSA (100 mM) (Promega), 1 µl of each primer (10 µM), 0.5 µl dNTPs (2.5 mM) and 0.25 µl Taq polymerase (5 U/µl) (Qiagen). The initial denaturation step of 3 minutes at 96 °C was followed by 40 cycles of 15 seconds at 96 °C, 30 seconds at 50 °C and 40 seconds at 72 °C, ending with a final extension step of 5 minutes at 72 °C. Amplified products were sequenced in both directions at Baseclear B.V. (Leiden, the Netherlands), combined and checked for errors in Geneious v8.1.9 (Kearse et al., 2012) and the final sequences were aligned using MAFFT v7.017 (Katoh and Standley, 2013).

Haplotype (H_d) and nucleotide diversity (π) were calculated using DnaSP v6.12.01 (Rozas et al., 2017) and reported for each station for both COI and 28S fragments. The 28S data were phased in DnaSP using PHASE (Stephens and Donnelly, 2003) with default settings of 100 iterations, thinning interval of 1 and burn-in of 100. Stations were tested for adherence to neutrality assumptions, with Tajima's D (Tajima, 1989) calculated in Arlequin version 3.5.2.2 (Excoffier and Lischer, 2010). Minimum-spanning networks for both genes were calculated and visualized in POPART (Leigh and Bryant, 2015). The phased allele alignment was used to construct the 28S haplotype network.

To identify population subdivision, genetic diversity between populations from different stations was quantified by pairwise Φ_{ST} (Holsinger and Weir, 2009), using the pairwise difference method in Arlequin with 10,000 permutations, and significant comparisons were identified after strict Bonferroni correction (Rice, 1989). Hierarchical Analysis of Molecular Variance (AMOVA) tests were conducted for both genes separately with sampling sites partitioned into (1) four 'ocean province' groups and (2) three abundance-based 'population' groups (Table 1). To test for isolation-by-distance (IBD), Mantel tests with 10,000 permutations were conducted in Arlequin to assess correlation between the COI or 28S pairwise Φ_{ST} matrices and the geographic distance matrix. Pairwise geographic distance between stations was calculated using the R package geodist with geodesic measures (Padgham and Sumner, 2020). The distance between stations 62 and 64 was averaged since they were analysed as a single site. The pairwise scatter plots were produced in R with ggplot2 (Wickham, 2016).

Morphometric analyses

Variation in shell shape was assessed using geometric morphometric measurements of 136 undamaged adult shells (shell length > 0.9mm). Specimens were positioned in a standardised apertural orientation and photographed using a Zeiss V20 stacking stereomicroscope with Axiovision software, prior to destructive DNA extraction. The resulting images were processed and digitised at eight (semi-)landmarks in TpsUtil and TpsDig (Rohlf, 2015) (Figure S2). The coordinates of the (semi-)landmarks were analysed in TpsRelW (Rohlf, 2015), using a generalised least-squares Procrustes superimposition (Rohlf and Slice, 1990; Zelditch et al., 2004). To ensure

that the digitisation process was repeatable, a subset of 30 individuals was photographed and digitised twice, with eight landmarks placed on each image. Centroid size and relative warp (RW) scores between the pairs of images per specimen after Procrustes Fit were compared using intra-class coefficient (ICC) in PAST3.0 (Hammer et al., 2001), and ICC values > 0.75 were considered sufficiently repeatable.

We tested for significant variation in shell shape across genetically distinct groups with a non-parametric Permutational Multivariate Analysis of Variance (one-way PERMANOVA) using Euclidean distances and 9999 permutations in R with vegan (Oksanen et al., 2019). Only centroid size and repeatable RWs were used in the one-way PERMANOVA, and strict Bonferroni corrections were applied for the pairwise PERMANOVAs. The first two RW axes were plotted to visualise shell shape variation for different groups of *L. bulimoides*. Additionally, a Canonical Variates Analysis (CVA) was conducted in R (R Core Team, 2017) to discriminate shell morphometric differences between groups. A one-way ANOVA with a post-hoc Tukey HSD test was also conducted in R to test if the means of the canonical variate for each group were different from the other groups.

Results

Genetic variability

The COI locus was highly polymorphic with global haplotype diversity (H_d) of 0.9998 (range: 0.9820-1.000) and nucleotide diversity (π) of 0.05449 (range: 0.0283-0.0402) across all sites (Table 2). In the COI minimum-spanning network, 349 unique haplotypes were observed across 356 individuals, and there was no shared central haplotype (Figure 2). Instead, Atlantic individuals were clustered into two main haplogroups that were separated by 15 substitutions (haplogroups 1 & 2). A third cluster was comprised of all Pacific individuals, which was separated from Atlantic haplogroup 1 by 62 substitutions. Population samples from stations 13, 19, 25, 43 and 45 showed deviations from neutrality based on Tajima's D (Table 2).

The 28S gene was less polymorphic than COI, with a global H_d of 0.6740 (range: 0.3740-0.8550) and a π of 0.0027 (range: 0.0006-0.0035) across all sites (Table 2). This reduced polymorphism is visible in the 28S minimum-spanning network (Figure 2), with only 63 haplotypes observed across 344 individuals. The three most abundant haplotypes comprised 369, 129 and 45 copies, respectively, and included representatives from both Atlantic and Pacific samples. The second most abundant haplotype was mainly comprised of individuals sampled from the equatorial upwelling region. Only station 55 showed a deviation from neutrality based on Tajima's D (Table 2).

Spatial structure

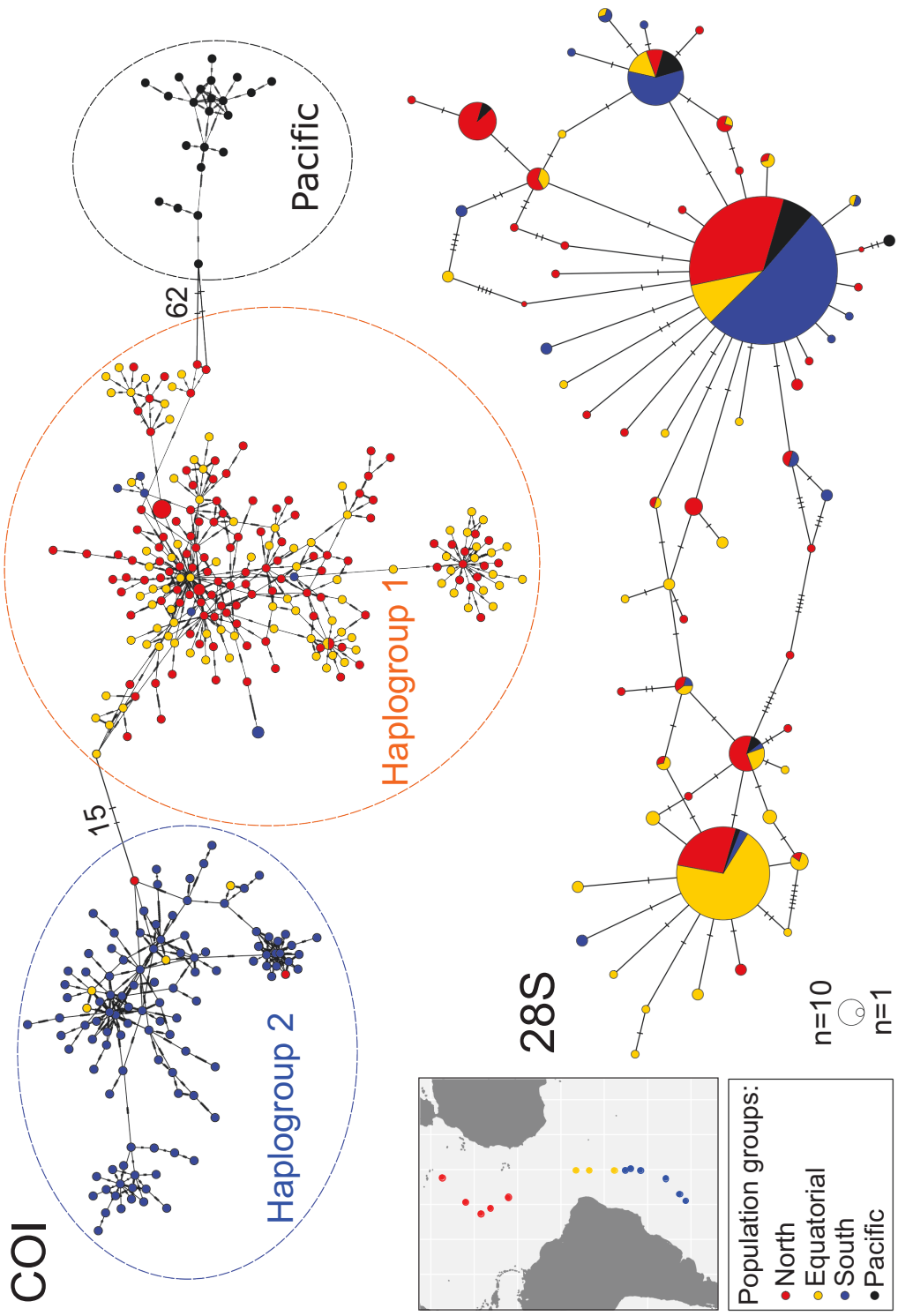
Overall Φ_{ST} calculated for both genes across the Atlantic supports the inference of significant spatial population structure (COI Φ_{ST} = 0.230, p = 0.00000; 28S Φ_{ST} = 0.255 p = 0.00000). Pairwise Φ_{ST} values for COI ranged from -0.017 to 0.451 for compar-

TABLE 2 Diversity indices of *Limacina bulimoides* for both mitochondrial cytochrome *c* oxidase I (COI) and nuclear 28S (28S) genes, including haplotype diversity (H_d), nucleotide diversity (π), Tajima's D (D). Significant Tajima's D values ($p < 0.05$), are indicated by *. Individuals from stations 62 and 64 were combined and analysed as one station due to the low numbers of individuals per site.

Station	Population groups	COI			28S		
		H_d	π	D	H_d	π	D
13	North	1	0.0283	-1.65*	0.623	0.0019	-1.12
19	North	1	0.0322	-1.47*	0.698	0.0028	0.359
23	North	1	0.0295	-1.44	0.855	0.0035	-0.517
25	North	0.982	0.0315	-1.57*	0.735	0.0024	-0.475
29	North	1	0.0299	-1.57	0.723	0.0031	1.46
43	Equatorial	1	0.0285	-1.53*	0.847	0.0036	-0.0628
45	Equatorial	1	0.0298	-1.46*	0.666	0.0024	-0.486
49	Equatorial	1	0.0379	-1.30	0.729	0.0032	-0.291
51	South	1	0.0356	-1.26	0.474	0.0011	-0.877
53	South	1	0.0372	-1.16	0.433	0.0006	-0.884
55	South	1	0.0402	-1.30	0.374	0.0007	-1.50*
60	South	1	0.0399	-1.19	0.458	0.0014	-1.10
62, 64	South	1	0.0382	-1.03	0.627	0.0010	-0.970
Pacific		1	0.0323	-1.33	0.818	0.0021	-0.728
Total		1	0.0545	-1.37	0.674	0.0027	-0.514

isons along the Atlantic transect and from 0.724 to 0.783 for comparisons between Atlantic and Pacific samples (TABLE 3). All comparisons between stations 13-49 (North & Equatorial) and stations 51-64 (South) were significant, with high Φ_{ST} values ranging from 0.250 to 0.451. We did not find any significant Φ_{ST} values in pairwise comparisons between stations 13-49. For nuclear 28S, pairwise Φ_{ST} values ranged from -0.001 to 0.621 for comparisons along the Atlantic transect, and from 0.079 to 0.532 for comparisons between Atlantic and Pacific samples (TABLE 3). The 28S pairwise Φ_{ST} values show a similar genetic break as for COI, separating stations 13-49 (North & Equatorial) from 51-64 (South), as 35 of the 40 pairwise Φ_{ST} values were significant (Φ_{ST} ranging from 0.012 to 0.621). There is additional spatial structuring separating samples from stations 13-29 (North) and 43-49 (Equatorial), as 13 out of 17 pairwise Φ_{ST} comparisons were significant (Φ_{ST} ranging from -0.001 to 0.497, TABLE 3).

Hierarchical analyses of molecular variance show that the spatial structure based on the trimodal pattern of abundance of *L. bulimoides* across a similar Atlantic transect (AMT24, FIGURE 3A) explains the distribution of genetic variation better than partitioning based on oceanographic parameters (TABLE 4). For both COI and 28S fragments, the fixation index among groups was higher when partitioned among three abundance-based 'population groups' (COI: $\Phi_{CT} = 0.285$, $p = 0.00129$; 28S: $\Phi_{CT} = 0.280$, $p = 0.00010$) than among four 'ocean provinces' (COI: $\Phi_{CT} = 0.205$, $p = 0.00446$; 28S: $\Phi_{CT} = 0.140$, $p = 0.0620$). Thus, we find a strong primary dispersal barrier separating stations 13-49 (North & Equatorial) from stations 51-64


COI
Haplogroup 2
Haplogroup 1
15
62
Pacific
28S
n=10
n=1
Population groups:
North
Equatorial
South
Pacific

(South) located at 15-18°S that is supported by both genes (Figure 3A, B). Surprisingly, this barrier is located in the core of the southern oligotrophic gyre and appears incongruent with any strong oceanographic gradients (Figure 3C, D). A more subtle secondary barrier is located in the region 15°N-4°S, which corresponds to the equatorial upwelling province and is mainly supported by 28S data.

We did not find obvious patterns of IBD across the Atlantic transect (Figure S3). Although there was a significant correlation between the pairwise Φ_{ST} and geographic distance for COI along the complete Atlantic transect (r = 0.520, p = 0.00120, Table S2), this relationship breaks down when population groups on either side of the primary barrier at 15-18°S were analysed separately (North + Equatorial: r = 0.117, p = 0.269; South: r = 0.0968, p= 0.418, Table S2). Pairwise Φ_{ST} values were uniformly high regardless of geographic distance when comparing stations on opposite sides of this dispersal barrier, and uniformly low when comparing stations within population groups (Figure S3A). This suggests that the significant correlation along the entire transect can be explained by the inclusion of high pairwise Φ_{ST} values between the stations across the barrier. For 28S, correlation of genetic and geographic distance matrices was not observed for the entire Atlantic transect (r = -0.0942, p = 0.770), but only for the North + Equatorial group (r = 0.759, p = 0.003, Table S2; Figure S3B). Again, this correlation is likely driven by the high Φ_{ST} values for comparisons between North and Equatorial samples on either side of the secondary dispersal barrier. No IBD patterns were found when any of the population groups were analysed separately.

The two mitochondrial haplogroups appear diagnostic for individuals on either side of the primary dispersal barrier with most individuals from the North and Equatorial stations belonging to COI haplogroup 1 and most individuals from the South belonging to COI haplogroup 2 (Figure 3B). However, we found 13 exceptions that are probably expatriates, or individuals sampled outside of their core range. Seven expatriates from the South clustered within haplogroup 1, and six expatriates from the North + Equatorial groups clustered within haplogroup 2 (Table S3; Figure 3B). We could verify the life stage for nine of these expatriates, of which three were adults (shell length >0.9 mm) and six were juveniles (Table S3). We assessed whether the 28S and mitochondrial haplotypes were congruent for 12 of these expatriates, but these results were inconclusive due to a lack of variability in the 28S fragment (Figure S4A).

Figure 2 Minimum spanning networks of *Limacina bulimoides* gene fragments of mitochondrial cytochrome *c* oxidase I (COI, top) and nuclear 28S rDNA (28S, bottom). The size of the filled circles represents the number of individuals with each haplotype (COI, 28S), with the smallest circles representing one individual with that haplotype, while colour represents abundance-based population groups (see map insert and Table 1). For the COI network, the dashed lines around the haplotype clusters indicate the three main haplogroups. Hatch marks on the branches represent the number of mutational steps, with large numbers of mutations between the haplogroups indicated with numbers.

TABLE 3 Pairwise Φ_{ST} for mitochondrial cytochrome *c* oxidase I (COI) (below diagonal) and nuclear 28S (28S) genes (above the diagonal) between all samples of *Limacina bulimoides*. Significant Φ_{ST} values after Bonferroni correction ($\alpha = 0.05$, $p < 0.000549$) are in bold.

	North					Equatorial			South					Pacific
Stations	13	19	23	25	29	43	45	49	51	53	55	60	62, 64	
13		**0.061**	0.039	0.034	**0.153**	**0.329**	**0.497**	**0.328**	0.025	**0.054**	**0.039**	**0.026**	0.037	**0.157**
19	0.010		0.005	0.039	0.020	**0.150**	**0.321**	**0.153**	**0.093**	**0.160**	**0.162**	**0.077**	**0.134**	**0.167**
23	0.000	-0.001		0.042	0.032	**0.163**	**0.329**	**0.168**	**0.070**	**0.120**	**0.134**	**0.062**	**0.101**	**0.154**
25	0.014	0.000	-0.004		**0.114**	**0.268**	**0.447**	**0.270**	0.042	**0.092**	**0.081**	0.035	0.062	**0.079**
29	0.012	0.014	0.003	-0.002		0.057	**0.205**	0.056	**0.190**	**0.273**	**0.304**	**0.137**	**0.241**	**0.244**
43	0.041	0.010	0.019	0.028	0.056		0.045	-0.001	**0.360**	**0.443**	**0.488**	**0.279**	**0.388**	**0.374**
45	0.002	0.000	0.006	0.003	0.005	0.034		0.052	**0.534**	**0.593**	**0.621**	**0.448**	**0.567**	**0.532**
49	0.019	0.011	0.012	0.006	0.022	0.032	0.012		**0.354**	**0.427**	**0.463**	**0.286**	**0.396**	**0.376**
51	**0.438**	**0.414**	**0.431**	**0.398**	**0.424**	**0.451**	**0.419**	**0.299**		0.023	0.013	0.012	0.065	**0.181**
53	**0.380**	**0.355**	**0.370**	**0.341**	**0.364**	**0.392**	**0.364**	**0.250**	0.007		0.036	0.023	0.070	**0.216**
55	**0.373**	**0.351**	**0.359**	**0.331**	**0.353**	**0.378**	**0.360**	**0.254**	0.005	-0.017		**0.028**	0.058	**0.242**
60	**0.408**	**0.389**	**0.400**	**0.369**	**0.391**	**0.416**	**0.398**	**0.289**	0.009	0.011	0.004		0.023	**0.106**
62, 64	**0.389**	**0.366**	**0.382**	**0.345**	**0.371**	**0.401**	**0.379**	**0.253**	0.022	-0.011	-0.009	0.021		**0.136**
Pacific	**0.780**	**0.769**	**0.781**	**0.773**	**0.778**	**0.783**	**0.772**	**0.742**	**0.759**	**0.748**	**0.731**	**0.724**	**0.735**	

Phenotypic variability

The centroid size and RW axes 1, 2, 4, 6, and 7 were repeatable and included in the final analyses. Based on the complete dataset of adult *L. bulimoides*, including Pacific individuals (N = 136), the repeatable RWs accounted for 85.21% of the total shell shape variation (Table S4). Shell shape variation for all specimens (N = 136) and Atlantic specimens only (N = 111) was visualised on the first two axes, which explain the most variation: RW1 (All: 61.08%; Atlantic: 62.41%) and RW2 (All: 13.45%;

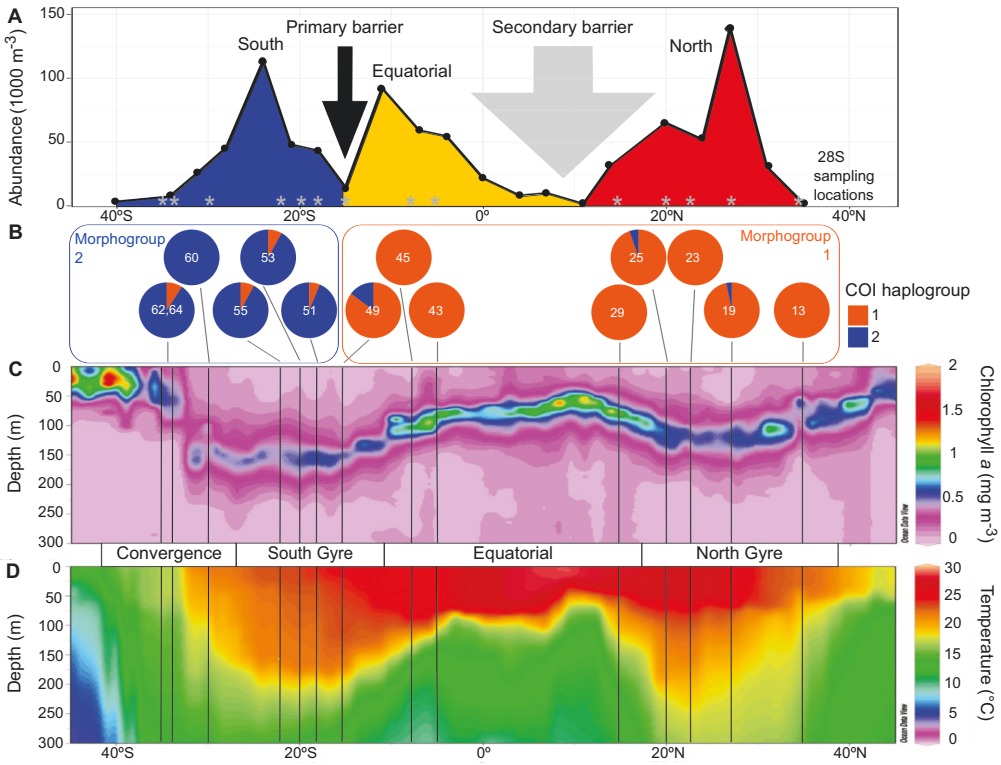


Figure 3 Environmental profile of two oceanic dispersal barriers for *Limacina bulimoides* in the Atlantic Ocean. The primary dispersal barrier is located at 15-18°S, while the secondary barrier is located in the region between 15°N and 4°S. (A) Abundance of *L. bulimoides* per 1000 m^3 seawater along the meridional transect sampled in 2014 (data from Burridge et al. 2017). Population groups are coloured as in Figure 2, supported by 28S genetic structuring (see Table 3) along the Atlantic Meridional Transect 22 (AMT22) sampled in 2012, locations indicated by grey asterisks. (B) Piecharts showing mitochondrial cytochrome *c* oxidase I (COI) haplogroup frequencies per sampling site with the division into two morphogroups superimposed (see also Figure 4B). (C) Ocean section plots of chlorophyll *a* concentration and (D) seawater temperature from the upper 300 meters of the water column with latitude. Sampling locations are marked by vertical lines and major oceanographic provinces (Convergence, South Gyre, Equatorial, and North Gyre) are indicated. Plots B, C and D derive from material sampled in 2012 (AMT22 transect).

TABLE 4 Results of hierarchical AMOVA for mitochondrial cytochrome *c* oxidase I (COI) (n = 334) and nuclear 28S (28S) (n = 324, 2n = 648) sequence variants in Atlantic individuals of *Limacina bulimoides*, partitioned based on four ocean provinces or three abundance-based population groups (TABLE 1). Significance levels for fixation indices are indicated as follows: n.s.: not significant, *: p < 0.05, **: p < 0.01, ***: p < 0.001.

Gene	Grouping	df	Variance components	% variation	Fixation indices
COI	Ocean provinces	3	Among groups	20.49	Φ_{CT} = 0.205**
		9	Among populations within groups	5.91	Φ_{SC} = 0.0743***
		321	Within populations	73.6	Φ_{ST} = 0.264***
	Population groups	2	Among groups	28.51	Φ_{CT} = 0.285**
		10	Among populations within groups	0.61	Φ_{SC} = 0.00849*
		321	Within populations	70.88	Φ_{ST} = 0.291***
28S	Ocean provinces	3	Among groups	14.02	Φ_{CT} = 0.140(n.s.)
		9	Among populations within groups	13.3	Φ_{SC} = 0.155***
		1283	Within populations	72.67	Φ_{ST} = 0.273***
	Population groups	2	Among groups	27.98	Φ_{CT} = 0.280***
		10	Among populations within groups	2.88	Φ_{SC} = 0.0400***
		1283	Within populations	69.13	Φ_{ST} = 0.309***

Atlantic: 13.43%) (FIGURE 4). The thin-plate spline reconstructions show that a positive RW1 is associated with a taller spire, while a negative RW1 is associated with a shorter, more rounded spire. Along RW2, larger values indicate wider shells.

We did not observe a significant difference in shell shape between Atlantic and Pacific *L. bulimoides* individuals (N = 136, F(1,135) = 1.66, p = 0.204). Shell shapes of the Pacific individuals (N = 25) fall within the variation observed for the Atlantic individuals (N = 111, FIGURE 4A). When Atlantic individuals were classified according to their spatial structure with North, Equatorial and South population groups, we found a significant difference in shell shape between groups based on the one-way PERMANOVA (F(2,110) = 14.53, p = 0.0001). Shell shape was different between the North and South (F = 35.9, p = 0.003), whereas comparisons of shell shape between the North and Equatorial or South and Equatorial groups were not significant (F = 2.74, p = 0.113 and F = 2.67, p = 0.116, respectively). This result could be due to the low sample size of the Equatorial group (N = 9). Individuals from the North population group generally had a higher, narrower spire compared to individuals from the South (FIGURE 4B). Similarly, there is a significant difference in shell shape across all three groups (F(2,108) = 29.79, p = 0.00000) based on the CVA. Significant shell shape differences were recorded for both North-South (p = 0.00000) and Equatorial-South (p = 0.00003) comparisons, but not between North and Equatorial groups (p = 0.889) based on Tukey HSD multiple pairwise comparisons (FIGURE S5, TABLE S5). Thus, the spatial structuring of samples in terms of shell shape matches the partitioning based on mitochondrial genetic variation (FIGURE 3B).

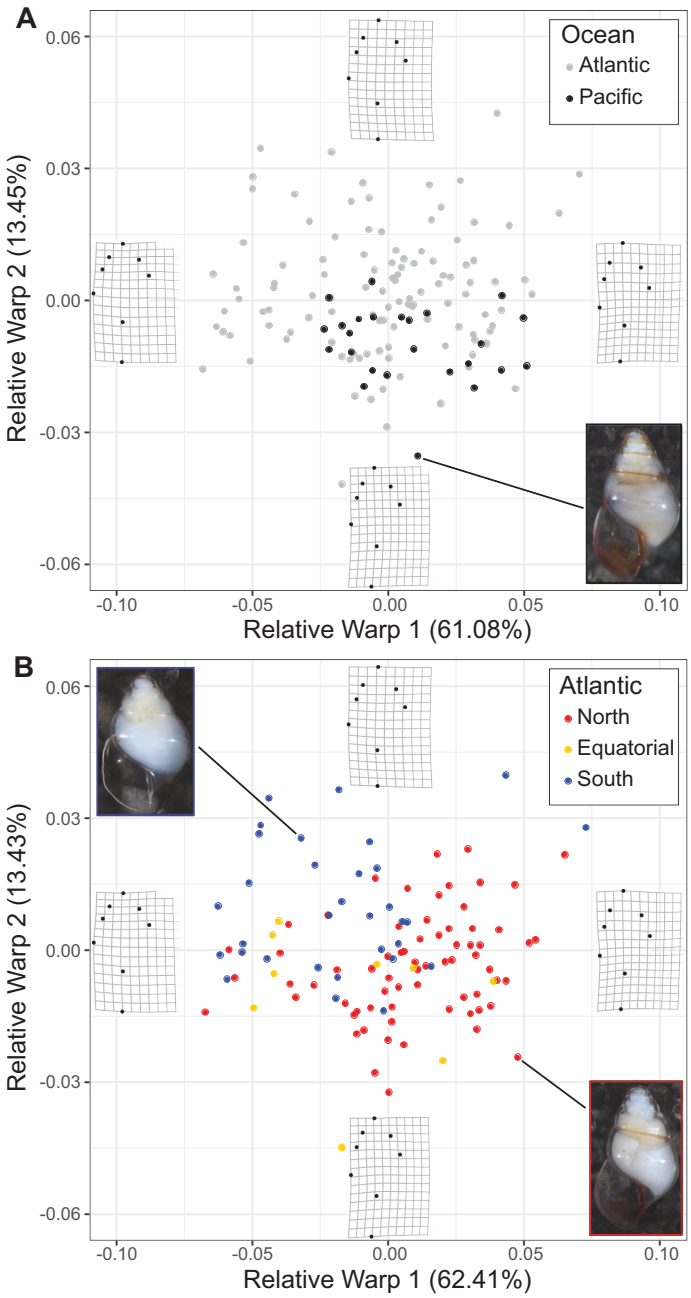


FIGURE 4 Shell shape variation in adult *Limacina bulimoides* visualised for the first two Relative Warp axes for A) Atlantic and Pacific individuals (N = 136), and B) Atlantic individuals (N = 111) coloured according to abundance-based population groups. Thin plate splines of the most positive and negative deformations along the axes are indicated to depict the variation in shell shape. Photographs of typical individuals from the Atlantic and Pacific (A) and Atlantic North and South population groups (B) are also shown.

Discussion

We found high levels of genetic variability and significant spatial population structure in the pteropod *Limacina bulimoides* along a latitudinal transect (35°N-36°S) in the Atlantic Ocean. Specifically, we have evidence for a primary dispersal barrier located in the southern subtropical gyre at 15-18°S and a more subtle, secondary barrier located in the region 15°N-4°S. This finding is in stark contrast to the historical perception of the pelagic environment as a homogeneous habitat lacking strong isolating barriers, and of holoplanktonic organisms as characterized by high dispersal and broad biogeographic ranges (Norris, 2000; Pierrot-Bults and van der Spoel, 1979). Because shelled pteropods are regarded as sentinel species to assess the impacts of ocean acidification (Bednaršek et al., 2017b), it is important to document their genetic variation and population structure, as these characteristics will influence their response to natural selection and global change (Barrett and Schluter, 2008; Donelson et al., 2019; Riebesell and Gattuso, 2015; Sanford and Kelly, 2011). This is the first study examining genetic structuring of a shelled pteropod with rigorous sampling at an ocean basin scale. The presence of strong oceanic dispersal barriers may indicate that the potential for range shifts and adaptive responses to a changing ocean could be more limited than expected, and that local adaptation should also be taken into account, for instance when modelling (future) species distribution patterns. Here, we discuss the nature and possible drivers of the oceanic dispersal barriers reported here for *L. bulimoides*, and more generally for pelagic taxa.

The primary dispersal barrier in the Atlantic Ocean probably represents a species barrier, across which populations on either side are reproductively isolated. This inference is supported by congruent differentiation at independently-inherited genetic loci (COI and 28S) and congruent genetic and shell shape differences. Despite the significant IBD test for the COI gene across the Atlantic transect, the disappearance of this result when populations across the primary barrier were tested separately supports the inference that these populations are spatially discrete evolutionary lineages (Teske et al., 2018). Nevertheless, a direct test of reproductive isolation is not possible as the high mortality of pteropods under laboratory conditions precludes meaningful crossing experiments (Howes et al., 2014). Genome-wide data would be necessary to determine the strength of reproductive isolation by quantifying divergence at many loci across the genome (Butlin and Stankowski, 2020; Gagnaire, 2020).

The observed morphometric variation was associated with the primary genetic barrier and not with any strong environmental transitions. Hence, it is more likely that the shell shape differences are the result of accumulated genetic differences than phenotypic plasticity resulting from distinct environments. Examples of strong plasticity in shells of marine molluscs are generally found across heterogeneous habitats, such as *Littorina saxatilis* with discrete 'crab' and 'wave' ecotypes (e.g., Hollander, Collyer, Adams, & Johannesson, 2006; Hollander and Butlin, 2010) or

Nacella concinna with continuous morphological differentiation across a depth gradient (Hoffman et al., 2010). Shell shape variation has also been shown to have a large genetic component, with heritabilities ranging between 0.36 and 0.71 in *Littorina saxatilis* (Conde-Padín et al., 2007) and between 0.16 and 0.56 in *Nucella lapillus* (Guerra-Varela et al., 2009). Furthermore, shell shape differences in conjunction with genetic differentiation have been used to identify species boundaries in different genera of holoplanktonic gastropods (Burridge, Goetze, Raes, Huisman, & Peijnenburg, 2015; Burridge et al., 2019; Wall-Palmer et al., 2018) using an integrative species concept (McManus and Katz, 2009; Padial et al., 2010). Hence, strong reproductive isolation is probably involved in the maintenance of this primary dispersal barrier in the southern Atlantic gyre.

The secondary dispersal barrier across the equatorial upwelling province is supported by limited evidence from the 28S gene (Table 3), and is congruent with known barriers in other pelagic taxa. The equatorial province has been suggested as an ecological barrier in some nekton species, such as Atlantic bluefin tuna (Briscoe et al., 2017) and whales (Holt et al., 2020), and as a barrier separating antitropical clades in planktonic foraminifers (Casteleyn et al., 2010). Dispersal barriers across the equatorial upwelling region have also been observed for other zooplankton. For example, in the pteropod genus *Cuvierina*, two distinct morphotypes are found in the Atlantic basin. One morphotype, 'atlantica', has disjunct populations in the northern and southern subtropical gyres, which are genetically differentiated, while the morphotype 'cancapae' dominates in equatorial waters (Burridge et al., 2015). Similar patterns of equatorial dispersal barriers were observed in the epipelagic gastropods *Protatlanta sculpta* and *P. souleyeti* (Wall-Palmer et al., 2016) and the mesopelagic copepod *Haloptilus longicornis* (Andrews et al., 2014; Norton and Goetze, 2013). We additionally observed a pattern of unique diversity in the equatorial region for the 28S gene fragment. This trend of higher equatorial diversity or equatorial endemics was also observed in populations of the copepods *Pleuromamma abdominalis* (Hirai et al., 2015) and *P. xiphias* (Goetze et al., 2017). It was suggested that this unique diversity resulted from a resident equatorial population, which was supplemented with expatriates by advection from the neighbouring gyre populations (Goetze et al., 2017). It is as yet unknown if expatriates in *L. bulimoides* survive to reproduce with the resident population and contribute to gene flow, or whether expatriates represent evolutionary dead-ends due to a mismatch of phenotype and environment (Marshall, Monro, Bode, Keough, & Swearer, 2010). Increased sampling intensity in the equatorial upwelling region, by having plankton tows with a greater amount of seawater filtered, would allow for the collection of more individuals in this region of low abundance for *L. bulimoides,* and a more thorough investigation into the prevalence and genetic identity of expatriates.

We identified two dispersal barriers for *L. bulimoides* in the Atlantic Ocean, of which one (the secondary barrier) appears more permeable than the other (the

primary barrier). The locations of these barriers coincide with regions of low abundance, congruent with the hypothesis that areas of less suitable habitat contribute to the maintenance of population structure in oceanic plankton (Goetze et al., 2017; Norris, 2000; Peijnenburg et al., 2004). The presence of expatriates suggests that Atlantic *L. bulimoides* populations could meet and potentially interbreed. This raises the question: what are the possible isolating mechanisms? These may include overlooked physical boundaries or ecological differences, including their depth habitat, feeding strategy or reproductive biology. Interestingly, the primary dispersal barrier is congruent with complex mesopelagic circulation patterns in the South Atlantic (~20°S, Sutton et al., 2017) where a shift in shrimp species (*Acanthephyra kingsleyi* to *A. quadrispinosa*) was observed at 18°S (Judkins, 2014). Little is known, however, about the behaviour or depth habitat of *L. bulimoides* across the Atlantic transect and thus it is possible that mesopelagic currents are important. Since internal fertilisation is the reproductive mode in *L. bulimoides*, assortative mate choice could also be a possible reproductive isolating mechanism, with chemical cues facilitating mate location and communication, as observed in a variety of marine gastropods (Ng et al., 2013). Other isolating mechanisms could involve differences in reproductive timing (McClary and Barker, 1998), ecological selection and community interactions (Whittaker and Rynearson, 2017), or post-zygotic barriers, such as reduced hybrid fitness (Lessios, 2007) or reduced hatching success of offspring (Ellingson and Krug, 2015). A broader range of ecological observations are required to determine the exact nature of dispersal barriers for *L. bulimoides*. For instance, future research to reveal population differences in feeding ecology or depth habitat could include the metabarcoding of gut contents, examination of the radula, or stable isotope analyses. Because we observe differences in the locations of maximum abundance and genetic transitions in different planktonic taxa sampled across similar Atlantic transects (Andrews et al., 2014; Goetze et al., 2015, 2017; Hirai et al., 2015; Norton and Goetze, 2013), we expect that dispersal barriers are species-specific and largely driven by the ecological characteristics of the respective species.

Many circumglobally distributed marine species have distinct populations or (sub)species in different ocean basins. These subdivisions generally result from divergence in allopatry, i.e., separated by current systems or continental land masses, and are often congruent with global biogeochemical provinces (Bowen et al., 2016; Costello et al., 2017; Reygondeau et al., 2012). The significant genetic divergence between Atlantic and Pacific *L. bulimoides* is in agreement with the patterns of divergence observed between ocean basins in other circumglobally distributed marine taxa, including fish (e.g., Graves and McDowell, 2015), benthic organisms (e.g., Fauvelot et al., 2020; Prazeres et al., 2020) and other zooplankton (reviewed in Peijnenburg and Goetze 2013). Based on a strict COI molecular clock rate of 2.4%/million years (MY) for gastropods (Hellberg and Vacquier, 1999), the divergence between the Atlantic and Pacific mitochondrial clades is estimated at

~4.6 MY, which roughly coincides with the emergence of the Isthmus of Panama (Bacon et al., 2015; O'Dea et al., 2016). The rise of the Isthmus of Panama has been directly linked to allopatric divergence across various marine taxa, due to the shoaling of the seaway, which reduced connectivity between mesopelagic populations at about 7 MY. This was followed by a decrease in seasonal upwelling due to the closing seaway at 4 MY, which cut off connectivity between plankton populations and decreased the input of primary productivity into the Caribbean (Knowlton, Weight, Solórzano, Mills, & Bermingham, 1993; Leigh, O'Dea, & Vermeij, 2014). This major marine biogeographic barrier was also used to calibrate the divergence between Atlantic and Pacific populations of three other pteropod species in a phylogenetic analysis (Burridge, Hörnlein et al., 2017).

The population structure within the Atlantic basin could be the result of different, not mutually exclusive, evolutionary scenarios. Dispersal barriers could have resulted from secondary contact following allopatric divergence in the past. The divergence time of the two Atlantic haplogroups across the primary barrier is estimated at 1.10 million years (MY) using a strict COI molecular clock (Hellberg and Vacquier, 1999). This estimate points to the potential effect of Pleistocene glacial cycles on population divergence of *L. bulimoides*. The effect of Pleistocene cycles as important drivers of phylogeographic structure has also been proposed for many other marine molluscs (e.g., Luttikhuizen, Drent, & Baker, 2003; Marko, 2004; Reid et al., 2006), including polar *Limacina* species (Sromek, Lasota, & Wolowicz, 2015). Another possible scenario is parapatric divergence, where semi-isolated populations living along an environmental cline diverged due to ecological selection. This would be a more plausible scenario for the secondary dispersal barrier across the equatorial upwelling gradients. The populations would progressively become more distinct, when the effect of divergent selection is greater than that of homogenising gene flow (Crow et al., 2007; Luttikhuizen et al., 2003; Postel et al., 2020; Schluter, 2009). In order to differentiate between historical divergence followed by secondary contact, ecological selection or a combination of the two, genome-wide analyses are required to investigate if genetic differentiation is randomly distributed across the genome (suggesting a gradual process) or associated with particular genes or environmental clines (indicative of ecological selection).

Holoplanktonic organisms are useful to gain insight into global biogeographical patterns in pelagic dispersal (Álvarez-Noriega et al., 2020; Bradbury et al., 2008), as their dispersal potential is not confounded by settlement processes that involve complex biophysical interactions with coastal or benthic habitats (Pineda et al., 2009; Prairie et al., 2012; Weersing and Toonen, 2009). To gain further insight into the nature and drivers of dispersal barriers in the open ocean, as well as to identify signals of selection across the genome, a broader range of ecological observations and in-depth analyses of genome-wide diversity in zooplankton will be necessary (Bucklin et al., 2018; Choo et al., 2020; Choquet et al., 2019; Gagnaire et al., 2015). This is important because marine zooplankton are key players in pelagic food webs and use-

ful indicators as rapid responders to environmental variation and climate change (Beaugrand et al., 2009). Shelled pteropods could be particularly suitable model organisms to gain insight into the evolutionary potential of marine zooplankton and to deepen our understanding of speciation processes in the open ocean for a number of reasons. First, there is growing interest in the group because of their vulnerability to ocean acidification and their potential use as bioindicators. Second, unlike most holoplanktonic animals, they have shells, which provide a record of the ecology and life history of an individual that can be easily measured. Third, pteropods are the only living planktonic animals with a good fossil record (Janssen and Peijnenburg, 2017; Peijnenburg et al., 2020), which allows for comparisons of morphological diversity in extant as well as extinct species and populations. We demonstrated significant spatial population structuring across and within ocean basins in *L. bulimoides*, a presumably circumglobal species. This underlying genetic variability and structure will influence the response of populations to global change, and in-depth study of their life-history traits, ecology and historical demography, in combination with experimental exposure to future conditions, will be required to more accurately predict their ability to adapt to a rapidly changing ocean.

Data availability

The nucleotide sequences for the COI and 28S genes are archived on NCBI GenBank with the accession numbers MN952611-MN952965 and MN950433-MN950775 respectively. Images for morphometric analyses are archived on Dryad (doi:10.5061/dryad.p8cz8w9nq).

Acknowledgements

This research was supported by the Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to K.T.C.A.P. Fieldwork also was supported by NSF grants OCE-1029478 and OCE-1338959 to E.G. We thank A. Burridge and M. Jungbluth for assistance at sea, and G. Tarran, A. Rees, and the captains and crew of the Atlantic Meridional Transect (AMT) cruises 22 and 24 for their leadership and assistance. We thank A. Tsuda for providing Pacific samples from the 2011/2012 R/V Hakuho-Maru KH-11-10 cruise. We are also grateful towards two anonymous reviewers whose comments have greatly improved the manuscript. The Atlantic Meridional Transect programme is funded by the UK Natural Environment Research Council through its National Capability Long-term Single Centre Science Programme, Climate Linked Atlantic Sector Science (grant number NE/R015953/1). This study contributes to the international IMBeR project and is contribution number 352 of the AMT programme.

Supplementary material

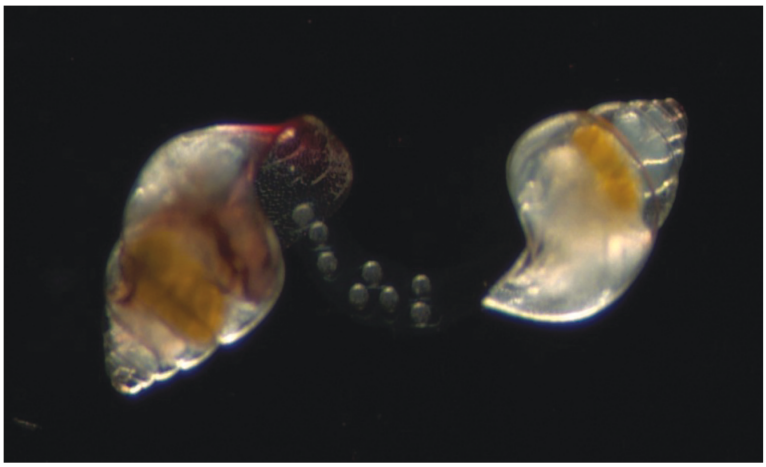


Figure S1 Two individuals of *Limacina. bulimoides*, with left individual laying free-floating egg strings, sampled from the North Atlantic (32°52.92 N, 26° 53.94 W) during the AMT27 cruise in 2017. See also Supplementary Video 1 which shows an individual of *L. bulimoides* (separate individual from this picture) in the process of laying egg strings.

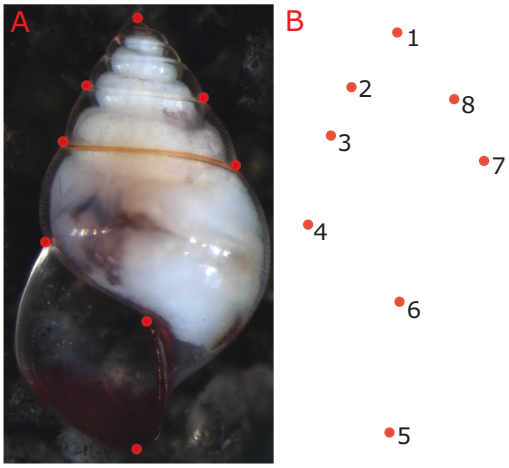


Figure S2 Representation of (semi-) landmark placement for 2D geometric morphometric analysis of shell shape for *Limacina bulimoides*. (A) The position of eight (semi-) landmarks on a shell. (B) Consensus plot of the shell shape calculated by the (semi-) landmark positions on 136 shells.

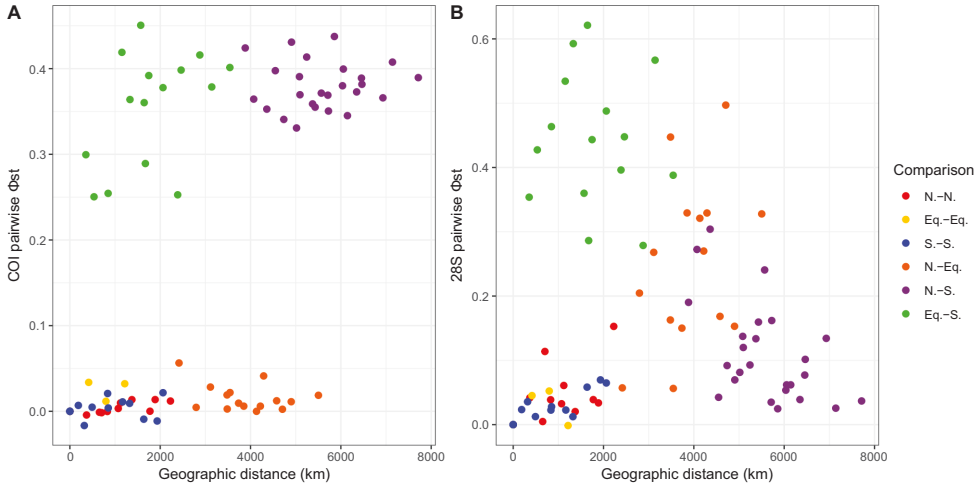


FIGURE S3 Scatter plots of genetic differentiation between Atlantic samples of *Limacina bulimoides* (pairwise Φ_{ST}) and geographic distance for (A) mitochondrial cytochrome *c* oxidase I (COI) and (B) nuclear 28S rDNA (28S). Comparisons within and between population groups are indicated with the following colour scheme: within North: red, within Equatorial: yellow, within South: blue, between North and Equatorial: orange, between North and South: purple, between Equatorial and South: green (see legend).

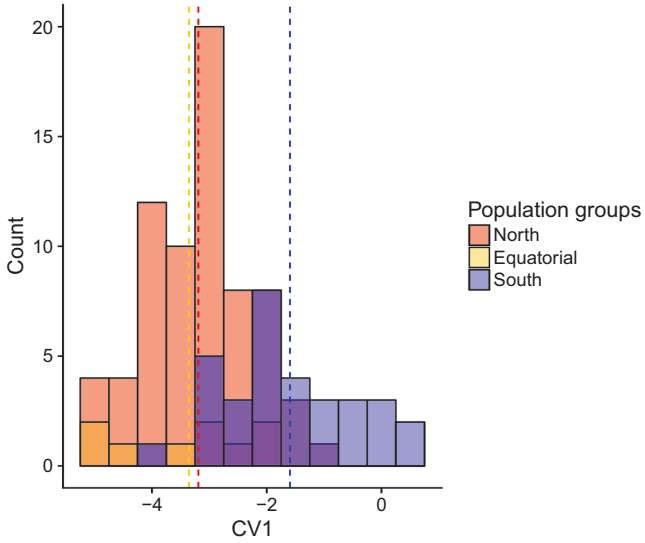


FIGURE S5 Canonical variate analysis of shell shape variation based on the repeatable Relative Warps (RWs), consisting of centroid size, RW1, RW2, RW4, RW6 and RW7. Shell shape variation between the Atlantic stations North (stations 13-29), Equatorial (43-45) and South (51-60) clades is maximised along the canonical variate 1 (CV1) axis. Group means for each group are indicated by dotted lines with the respective colour on the CV1 axis. Results of multiple pairwise comparisons of the means are reported in TABLE S5.

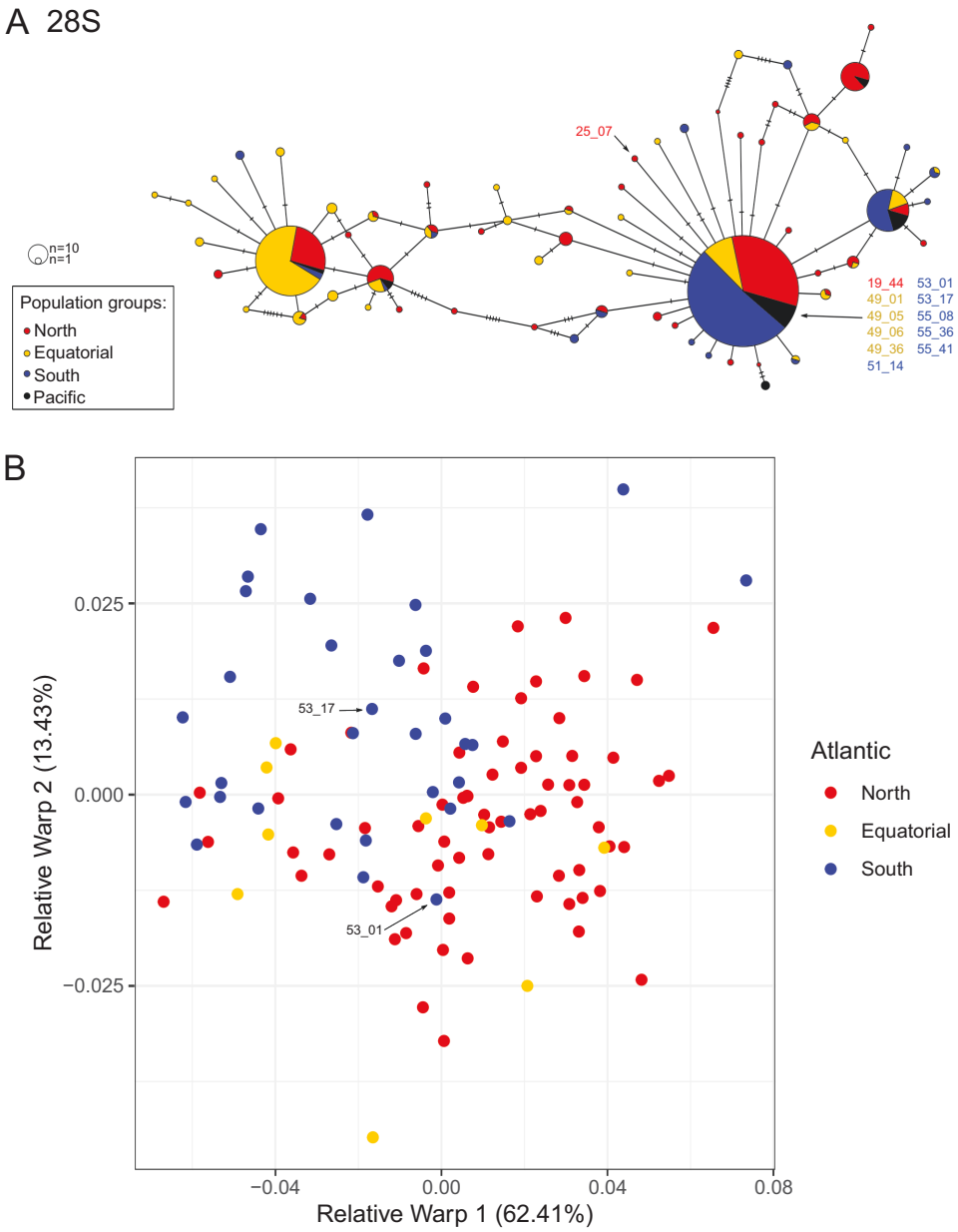


FIGURE S4 Information about the nuclear 28S haplotype (A) and morphotype (B) of the expatriates identified based on mitochondrial haplogroups (see also TABLE S3). 28S haplotype data was available for 12 individuals (A) but only two individuals were included in the morphometric analysis (B).

Table S1 Specimen information for the 413 *Limacina bulimoides* individuals included in this study, including specimen voucher, specimen ID, location, collection date, NCBI accession numbers for mitochondrial cytochrome *c* oxidase I (COI) and nuclear 28S rDNA (28S) genes, and if it was included in the morphometrics dataset.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342218	Lbul_AMT22_13_01	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952611	MN950433	
RMNH.MOL.342219	Lbul_AMT22_13_02	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952612	MN950434	
RMNH.MOL.342220	Lbul_AMT22_13_03	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952613	MN950435	
RMNH.MOL.342221	Lbul_AMT22_13_04	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952614	MN950436	
RMNH.MOL.342222	Lbul_AMT22_13_05	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952615	MN950437	
RMNH.MOL.342223	Lbul_AMT22_13_09	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952616	MN950438	
RMNH.MOL.342224	Lbul_AMT22_13_10	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952617	MN950439	
RMNH.MOL.342225	Lbul_AMT22_13_14	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952618	MN950440	
RMNH.MOL.342226	Lbul_AMT22_13_15	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952619	MN950441	
RMNH.MOL.342227	Lbul_AMT22_13_16	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952620	MN950442	
RMNH.MOL.342228	Lbul_AMT22_13_20	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952621	MN950443	
RMNH.MOL.342229	Lbul_AMT22_13_21	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952622	MN950444	
RMNH.MOL.342230	Lbul_AMT22_13_23	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952623	MN950445	
RMNH.MOL.342231	Lbul_AMT22_13_24	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952624	MN950446	
RMNH.MOL.342232	Lbul_AMT22_13_25	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952625	MN950447	
RMNH.MOL.342233	Lbul_AMT22_13_26	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952626	MN950448	
RMNH.MOL.342234	Lbul_AMT22_13_28	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952627	MN950449	Yes
RMNH.MOL.342235	Lbul_AMT22_13_29	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952628	MN950450	Yes
RMNH.MOL.342236	Lbul_AMT22_13_30	Atlantic	AMT22_13	34.37	-27.63	18/10/2012			Yes
RMNH.MOL.342237	Lbul_AMT22_13_31	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952629	MN950451	Yes
RMNH.MOL.342238	Lbul_AMT22_13_32	Atlantic	AMT22_13	34.37	-27.63	18/10/2012			Yes
RMNH.MOL.342239	Lbul_AMT22_13_33	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952630	MN950452	
RMNH.MOL.342240	Lbul_AMT22_13_34	Atlantic	AMT22_13	34.37	-27.63	18/10/2012			Yes
RMNH.MOL.342241	Lbul_AMT22_13_35	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952631	MN950453	Yes
RMNH.MOL.342242	Lbul_AMT22_13_36	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952632	MN950454	
RMNH.MOL.342243	Lbul_AMT22_13_37	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952633	MN950455	
RMNH.MOL.342244	Lbul_AMT22_13_38	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952634	MN950735	Yes
RMNH.MOL.342245	Lbul_AMT22_13_39	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952635	MN950456	Yes
RMNH.MOL.342246	Lbul_AMT22_13_40	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952636	MN950736	Yes
RMNH.MOL.342247	Lbul_AMT22_13_41	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952637	MN950737	
RMNH.MOL.342248	Lbul_AMT22_13_42	Atlantic	AMT22_13	34.37	-27.63	18/10/2012	MN952638	MN950738	
RMNH.MOL.342249	Lbul_AMT22_19_03	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342250	Lbul_AMT22_19_06	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952639	MN950457	
RMNH.MOL.342251	Lbul_AMT22_19_09	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952640	MN950458	Yes
RMNH.MOL.342252	Lbul_AMT22_19_10	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952641	MN950459	
RMNH.MOL.342253	Lbul_AMT22_19_11	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952642	MN950460	
RMNH.MOL.342254	Lbul_AMT22_19_12	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952643	MN950461	
RMNH.MOL.342255	Lbul_AMT22_19_13	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952644	MN950462	
RMNH.MOL.342256	Lbul_AMT22_19_14	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952645	MN950463	
RMNH.MOL.342257	Lbul_AMT22_19_15	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952646	MN950464	
RMNH.MOL.342258	Lbul_AMT22_19_16	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952647	MN950465	
RMNH.MOL.342259	Lbul_AMT22_19_17	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952648	MN950466	
RMNH.MOL.342260	Lbul_AMT22_19_18	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952649	MN950467	
RMNH.MOL.342261	Lbul_AMT22_19_19	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952650	MN950468	
RMNH.MOL.342262	Lbul_AMT22_19_20	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952651		
RMNH.MOL.342263	Lbul_AMT22_19_21	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952652	MN950469	
RMNH.MOL.342264	Lbul_AMT22_19_22	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952653	MN950470	
RMNH.MOL.342265	Lbul_AMT22_19_23	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952654	MN950471	
RMNH.MOL.342266	Lbul_AMT22_19_24	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952655	MN950472	
RMNH.MOL.342267	Lbul_AMT22_19_25	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952656	MN950473	
RMNH.MOL.342268	Lbul_AMT22_19_26	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952657	MN950474	
RMNH.MOL.342269	Lbul_AMT22_19_27	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952658	MN950475	
RMNH.MOL.342270	Lbul_AMT22_19_28	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342271	Lbul_AMT22_19_29	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342272	Lbul_AMT22_19_30	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342273	Lbul_AMT22_19_31	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952659	MN950476	
RMNH.MOL.342274	Lbul_AMT22_19_32	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342275	Lbul_AMT22_19_33	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952660	MN950477	Yes
RMNH.MOL.342276	Lbul_AMT22_19_34	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952661	MN950478	Yes
RMNH.MOL.342277	Lbul_AMT22_19_35	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952662		Yes
RMNH.MOL.342278	Lbul_AMT22_19_36	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952663	MN950479	Yes
RMNH.MOL.342279	Lbul_AMT22_19_37	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342280	Lbul_AMT22_19_38	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952664		Yes
RMNH.MOL.342281	Lbul_AMT22_19_39	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952665		Yes
RMNH.MOL.342282	Lbul_AMT22_19_40	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952666		Yes
RMNH.MOL.342283	Lbul_AMT22_19_41	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952667	MN950480	Yes

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342284	Lbul_AMT22_19_42	Atlantic	AMT22_19	27.60	-36.37	21/10/2012			Yes
RMNH.MOL.342285	Lbul_AMT22_19_44	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952668	MN950481	
RMNH.MOL.342286	Lbul_AMT22_19_45	Atlantic	AMT22_19	27.60	-36.37	21/10/2012	MN952669	MN950482	Yes
RMNH.MOL.342287	Lbul_AMT22_23_01	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952670		
RMNH.MOL.342288	Lbul_AMT22_23_02	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952671	MN950483	
RMNH.MOL.342289	Lbul_AMT22_23_03	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952672	MN950484	
RMNH.MOL.342290	Lbul_AMT22_23_04	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952673	MN950485	
RMNH.MOL.342291	Lbul_AMT22_23_06	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952674		
RMNH.MOL.342292	Lbul_AMT22_23_07	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952675	MN950486	
RMNH.MOL.342293	Lbul_AMT22_23_08	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952676	MN950487	
RMNH.MOL.342294	Lbul_AMT22_23_09	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952677	MN950488	
RMNH.MOL.342295	Lbul_AMT22_23_10	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952678	MN950489	
RMNH.MOL.342296	Lbul_AMT22_23_11	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952679	MN950490	
RMNH.MOL.342297	Lbul_AMT22_23_12	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952680	MN950491	
RMNH.MOL.342298	Lbul_AMT22_23_13	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952681	MN950492	
RMNH.MOL.342299	Lbul_AMT22_23_14	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952682	MN950493	
RMNH.MOL.342300	Lbul_AMT22_23_17	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952683	MN950494	
RMNH.MOL.342301	Lbul_AMT22_23_18	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952684	MN950495	
RMNH.MOL.342302	Lbul_AMT22_23_19	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952685	MN950496	
RMNH.MOL.342303	Lbul_AMT22_23_21	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952686	MN950497	
RMNH.MOL.342304	Lbul_AMT22_23_22	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952687	MN950498	
RMNH.MOL.342305	Lbul_AMT22_23_23	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952688		
RMNH.MOL.342306	Lbul_AMT22_23_24	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952689	MN950499	
RMNH.MOL.342307	Lbul_AMT22_23_25	Atlantic	AMT22_23	23.15	-40.63	23/10/2012			Yes
RMNH.MOL.342308	Lbul_AMT22_23_26	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952690	MN950500	Yes
RMNH.MOL.342309	Lbul_AMT22_23_27	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952691	MN950501	
RMNH.MOL.342310	Lbul_AMT22_23_28	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952692	MN950502	
RMNH.MOL.342311	Lbul_AMT22_23_29	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952693	MN950503	
RMNH.MOL.342312	Lbul_AMT22_23_30	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952694		
RMNH.MOL.342313	Lbul_AMT22_23_31	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952695	MN950504	
RMNH.MOL.342314	Lbul_AMT22_23_32	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952696	MN950505	
RMNH.MOL.342315	Lbul_AMT22_23_34	Atlantic	AMT22_23	23.15	-40.63	23/10/2012	MN952697	MN950506	
RMNH.MOL.342316	Lbul_AMT22_25_02	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952698	MN950507	Yes
RMNH.MOL.342317	Lbul_AMT22_25_03	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952699	MN950508	Yes

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342318	Lbul_AMT22_25_04	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952700	MN950509	Yes
RMNH.MOL.342319	Lbul_AMT22_25_05	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952701	MN950510	
RMNH.MOL.342320	Lbul_AMT22_25_06	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952702	MN950511	
RMNH.MOL.342321	Lbul_AMT22_25_07	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952703	MN950739	
RMNH.MOL.342322	Lbul_AMT22_25_08	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952704	MN950512	Yes
RMNH.MOL.342323	Lbul_AMT22_25_09	Atlantic	AMT22_25	20.40	-38.62	24/10/2012		MN950740	
RMNH.MOL.342324	Lbul_AMT22_25_10	Atlantic	AMT22_25	20.40	-38.62	24/10/2012			Yes
RMNH.MOL.342325	Lbul_AMT22_25_13	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952705	MN950513	Yes
RMNH.MOL.342326	Lbul_AMT22_25_14	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952706	MN950514	Yes
RMNH.MOL.342327	Lbul_AMT22_25_15	Atlantic	AMT22_25	20.40	-38.62	24/10/2012			Yes
RMNH.MOL.342328	Lbul_AMT22_25_16	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952707	MN950741	Yes
RMNH.MOL.342329	Lbul_AMT22_25_17	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952708	MN950515	Yes
RMNH.MOL.342330	Lbul_AMT22_25_18	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952709	MN950516	Yes
RMNH.MOL.342331	Lbul_AMT22_25_19	Atlantic	AMT22_25	20.40	-38.62	24/10/2012			Yes
RMNH.MOL.342332	Lbul_AMT22_25_20	Atlantic	AMT22_25	20.40	-38.62	24/10/2012			Yes
RMNH.MOL.342333	Lbul_AMT22_25_21	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952710	MN950517	Yes
RMNH.MOL.342334	Lbul_AMT22_25_22	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952711	MN950518	Yes
RMNH.MOL.342335	Lbul_AMT22_25_23	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952712		Yes
RMNH.MOL.342336	Lbul_AMT22_25_24	Atlantic	AMT22_25	20.40	-38.62	24/10/2012			Yes
RMNH.MOL.342337	Lbul_AMT22_25_25	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952713	MN950519	Yes
RMNH.MOL.342338	Lbul_AMT22_25_26	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952714	MN950520	Yes
RMNH.MOL.342339	Lbul_AMT22_25_27	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952715	MN950521	Yes
RMNH.MOL.342340	Lbul_AMT22_25_28	Atlantic	AMT22_25	20.40	-38.62	24/10/2012	MN952716	MN950522	Yes
RMNH.MOL.342341	Lbul_AMT22_29A_01	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952717	MN950523	Yes
RMNH.MOL.342342	Lbul_AMT22_29A_02	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342343	Lbul_AMT22_29A_03	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952718	MN950524	
RMNH.MOL.342344	Lbul_AMT22_29A_04	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952719	MN950525	
RMNH.MOL.342345	Lbul_AMT22_29A_05	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952720	MN950526	Yes
RMNH.MOL.342346	Lbul_AMT22_29A_06	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952721	MN950527	
RMNH.MOL.342347	Lbul_AMT22_29A_07	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952722	MN950528	Yes
RMNH.MOL.342348	Lbul_AMT22_29A_08	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952723	MN950529	
RMNH.MOL.342349	Lbul_AMT22_29A_09	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952724	MN950530	
RMNH.MOL.342350	Lbul_AMT22_29A_10	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952725	MN950531	
RMNH.MOL.342351	Lbul_AMT22_29A_11	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952726	MN950532	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342352	Lbul_AMT22_29A_14	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952727	MN950533	
RMNH.MOL.342353	Lbul_AMT22_29A_15	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952728	MN950534	
RMNH.MOL.342354	Lbul_AMT22_29A_17	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952729	MN950535	
RMNH.MOL.342355	Lbul_AMT22_29A_19	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952730	MN950536	
RMNH.MOL.342356	Lbul_AMT22_29A_21	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952731		
RMNH.MOL.342357	Lbul_AMT22_29A_22	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952732	MN950537	
RMNH.MOL.342358	Lbul_AMT22_29A_23	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952733	MN950538	
RMNH.MOL.342359	Lbul_AMT22_29A_24	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952734		
RMNH.MOL.342360	Lbul_AMT22_29A_25	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342361	Lbul_AMT22_29A_26	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342362	Lbul_AMT22_29A_27	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342363	Lbul_AMT22_29A_28	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342364	Lbul_AMT22_29A_29	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342365	Lbul_AMT22_29A_30	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952735	MN950539	Yes
RMNH.MOL.342366	Lbul_AMT22_29A_31	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342367	Lbul_AMT22_29A_32	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342368	Lbul_AMT22_29A_33	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342369	Lbul_AMT22_29A_34	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342370	Lbul_AMT22_29A_35	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342371	Lbul_AMT22_29A_36	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342372	Lbul_AMT22_29A_37	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952736	MN950540	Yes
RMNH.MOL.342373	Lbul_AMT22_29A_38	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012	MN952737	MN950541	Yes
RMNH.MOL.342374	Lbul_AMT22_29A_39	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342375	Lbul_AMT22_29A_40	Atlantic	AMT22_29A	15.30	-34.67	26/10/2012			Yes
RMNH.MOL.342376	Lbul_AMT22_43A_01	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952738	MN950542	
RMNH.MOL.342377	Lbul_AMT22_43A_03	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952739	MN950543	
RMNH.MOL.342378	Lbul_AMT22_43A_04	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952740	MN950544	
RMNH.MOL.342379	Lbul_AMT22_43A_05	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952741		
RMNH.MOL.342380	Lbul_AMT22_43A_06	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952742	MN950545	
RMNH.MOL.342381	Lbul_AMT22_43A_08	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952743	MN950546	
RMNH.MOL.342382	Lbul_AMT22_43A_11	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952744	MN950547	
RMNH.MOL.342383	Lbul_AMT22_43A_13	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952745	MN950548	
RMNH.MOL.342384	Lbul_AMT22_43A_20	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952746		
RMNH.MOL.342385	Lbul_AMT22_43A_21	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952747	MN950549	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342386	Lbul_AMT22_43A_22	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952748	MN950550	
RMNH.MOL.342387	Lbul_AMT22_43A_24	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952749		
RMNH.MOL.342388	Lbul_AMT22_43A_25	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952750	MN950551	
RMNH.MOL.342389	Lbul_AMT22_43A_26	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952751	MN950552	
RMNH.MOL.342390	Lbul_AMT22_43A_27	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012			Yes
RMNH.MOL.342391	Lbul_AMT22_43A_28	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952752		
RMNH.MOL.342392	Lbul_AMT22_43A_29	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952753		
RMNH.MOL.342393	Lbul_AMT22_43A_30	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952754	MN950553	
RMNH.MOL.342394	Lbul_AMT22_43A_31	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952755	MN950554	
RMNH.MOL.342395	Lbul_AMT22_43A_32	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952756	MN950555	
RMNH.MOL.342396	Lbul_AMT22_43A_33	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012			Yes
RMNH.MOL.342397	Lbul_AMT22_43A_34	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952757	MN950556	Yes
RMNH.MOL.342398	Lbul_AMT22_43A_35	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952758	MN950742	
RMNH.MOL.342399	Lbul_AMT22_43A_36	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952759	MN950557	
RMNH.MOL.342400	Lbul_AMT22_43A_37	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952760	MN950558	
RMNH.MOL.342401	Lbul_AMT22_43A_38	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952761	MN950743	
RMNH.MOL.342402	Lbul_AMT22_43A_39	Atlantic	AMT22_43A	-4.32	-25.02	1/11/2012	MN952762	MN950744	
RMNH.MOL.342403	Lbul_AMT22_45_01	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952763	MN950559	
RMNH.MOL.342404	Lbul_AMT22_45_02	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952764	MN950560	
RMNH.MOL.342405	Lbul_AMT22_45_03	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952765	MN950561	
RMNH.MOL.342406	Lbul_AMT22_45_04	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952766	MN950562	
RMNH.MOL.342407	Lbul_AMT22_45_05	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952767	MN950563	
RMNH.MOL.342408	Lbul_AMT22_45_06	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952768	MN950564	
RMNH.MOL.342409	Lbul_AMT22_45_08	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952769	MN950565	
RMNH.MOL.342410	Lbul_AMT22_45_09	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952770	MN950566	
RMNH.MOL.342411	Lbul_AMT22_45_10	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952771	MN950567	
RMNH.MOL.342412	Lbul_AMT22_45_11	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952772	MN950568	
RMNH.MOL.342413	Lbul_AMT22_45_12	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952773	MN950569	
RMNH.MOL.342414	Lbul_AMT22_45_13	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952774	MN950570	
RMNH.MOL.342415	Lbul_AMT22_45_14	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952775	MN950571	
RMNH.MOL.342416	Lbul_AMT22_45_15	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952776	MN950572	
RMNH.MOL.342417	Lbul_AMT22_45_16	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952777	MN950573	
RMNH.MOL.342418	Lbul_AMT22_45_17	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952778	MN950574	
RMNH.MOL.342419	Lbul_AMT22_45_18	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952779	MN950575	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342420	Lbul_AMT22_45_19	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952780	MN950576	
RMNH.MOL.342421	Lbul_AMT22_45_20	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952781	MN950577	
RMNH.MOL.342422	Lbul_AMT22_45_21	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952782	MN950578	
RMNH.MOL.342423	Lbul_AMT22_45_22	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952783	MN950579	
RMNH.MOL.342424	Lbul_AMT22_45_23	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012			Yes
RMNH.MOL.342425	Lbul_AMT22_45_24	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952784	MN950580	Yes
RMNH.MOL.342426	Lbul_AMT22_45_25	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952785	MN950581	
RMNH.MOL.342427	Lbul_AMT22_45_26	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952786	MN950582	Yes
RMNH.MOL.342428	Lbul_AMT22_45_27	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952787	MN950583	Yes
RMNH.MOL.342429	Lbul_AMT22_45_28	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952788	MN950584	Yes
RMNH.MOL.342430	Lbul_AMT22_45_29	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952789	MN950585	
RMNH.MOL.342431	Lbul_AMT22_45_30	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952790	MN950586	Yes
RMNH.MOL.342432	Lbul_AMT22_45_31	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952791	MN950587	
RMNH.MOL.342433	Lbul_AMT22_45_32	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952792		
RMNH.MOL.342434	Lbul_AMT22_45_33	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952793	MN950745	
RMNH.MOL.342435	Lbul_AMT22_45_34	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952794	MN950746	
RMNH.MOL.342436	Lbul_AMT22_45_35	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952795	MN950747	
RMNH.MOL.342437	Lbul_AMT22_45_38	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012		MN950748	
RMNH.MOL.342438	Lbul_AMT22_45_39	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952796	MN950749	
RMNH.MOL.342439	Lbul_AMT22_45_40	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952797	MN950588	
RMNH.MOL.342440	Lbul_AMT22_45_41	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952798	MN950589	
RMNH.MOL.342441	Lbul_AMT22_45_42	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952799	MN950590	
RMNH.MOL.342442	Lbul_AMT22_45_43	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952800	MN950591	
RMNH.MOL.342443	Lbul_AMT22_45_44	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952801	MN950592	
RMNH.MOL.342444	Lbul_AMT22_45_45	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952802	MN950593	
RMNH.MOL.342445	Lbul_AMT22_45_46	Atlantic	AMT22_45	-8.08	-25.03	3/11/2012	MN952803	MN950594	
RMNH.MOL.342446	Lbul_AMT22_49_01	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952804	MN950595	
RMNH.MOL.342447	Lbul_AMT22_49_04	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952805		
RMNH.MOL.342448	Lbul_AMT22_49_05	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952806	MN950596	
RMNH.MOL.342449	Lbul_AMT22_49_06	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952807	MN950597	
RMNH.MOL.342450	Lbul_AMT22_49_07	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952808	MN950598	
RMNH.MOL.342451	Lbul_AMT22_49_08	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952809	MN950599	
RMNH.MOL.342452	Lbul_AMT22_49_09	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952810	MN950600	
RMNH.MOL.342453	Lbul_AMT22_49_10	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952811	MN950601	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342454	Lbul_AMT22_49_11	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952812	MN950602	
RMNH.MOL.342455	Lbul_AMT22_49_12	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952813	MN950603	
RMNH.MOL.342456	Lbul_AMT22_49_13	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952814	MN950604	
RMNH.MOL.342457	Lbul_AMT22_49_14	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952815	MN950605	
RMNH.MOL.342458	Lbul_AMT22_49_15	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952816	MN950606	
RMNH.MOL.342459	Lbul_AMT22_49_16	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952817	MN950607	
RMNH.MOL.342460	Lbul_AMT22_49_17	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952818	MN950608	
RMNH.MOL.342461	Lbul_AMT22_49_18	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952819	MN950609	
RMNH.MOL.342462	Lbul_AMT22_49_19	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952820	MN950610	
RMNH.MOL.342463	Lbul_AMT22_49_20	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952821	MN950611	
RMNH.MOL.342464	Lbul_AMT22_49_21	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952822	MN950612	
RMNH.MOL.342465	Lbul_AMT22_49_22	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952823	MN950613	
RMNH.MOL.342466	Lbul_AMT22_49_23	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952824	MN950614	
RMNH.MOL.342467	Lbul_AMT22_49_24	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952825	MN950615	
RMNH.MOL.342468	Lbul_AMT22_49_25	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952826	MN950616	
RMNH.MOL.342469	Lbul_AMT22_49_26	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952827	MN950617	
RMNH.MOL.342470	Lbul_AMT22_49_28	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952828	MN950750	
RMNH.MOL.342471	Lbul_AMT22_49_29	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012		MN950751	
RMNH.MOL.342472	Lbul_AMT22_49_30	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012		MN950752	
RMNH.MOL.342473	Lbul_AMT22_49_31	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952829	MN950753	
RMNH.MOL.342474	Lbul_AMT22_49_33	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012		MN950754	
RMNH.MOL.342475	Lbul_AMT22_49_34	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012		MN950755	
RMNH.MOL.340293	Lbul_AMT22_49_35	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012		MK635470	
RMNH.MOL.342477	Lbul_AMT22_49_36	Atlantic	AMT22_49	-15.30	-25.07	5/11/2012	MN952830	MN950618	
RMNH.MOL.342478	Lbul_AMT22_51_02	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952831	MN950619	
RMNH.MOL.342479	Lbul_AMT22_51_03	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952832	MN950620	
RMNH.MOL.342480	Lbul_AMT22_51_04	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952833	MN950621	
RMNH.MOL.342481	Lbul_AMT22_51_06	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952834	MN950622	
RMNH.MOL.342482	Lbul_AMT22_51_11	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952835	MN950623	Yes
RMNH.MOL.342483	Lbul_AMT22_51_12	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952836	MN950624	
RMNH.MOL.342484	Lbul_AMT22_51_13	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952837	MN950625	
RMNH.MOL.342485	Lbul_AMT22_51_14	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952838	MN950626	
RMNH.MOL.342486	Lbul_AMT22_51_15	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952839	MN950627	
RMNH.MOL.342487	Lbul_AMT22_51_16	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952840	MN950628	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342488	Lbul_AMT22_51_18	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952841	MN950629	
RMNH.MOL.342489	Lbul_AMT22_51_19	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952842	MN950630	
RMNH.MOL.342490	Lbul_AMT22_51_20	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952843	MN950631	
RMNH.MOL.342491	Lbul_AMT22_51_21	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952844	MN950632	
RMNH.MOL.342492	Lbul_AMT22_51_22	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952845	MN950633	
RMNH.MOL.342493	Lbul_AMT22_51_24	Atlantic	AMT22_51	-18.50	-25.10	6/11/2012	MN952846	MN950634	
RMNH.MOL.342494	Lbul_AMT22_53_01	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952847	MN950635	Yes
RMNH.MOL.342495	Lbul_AMT22_53_02	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952848	MN950636	Yes
RMNH.MOL.342496	Lbul_AMT22_53_03	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952849	MN950637	Yes
RMNH.MOL.342497	Lbul_AMT22_53_04	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012			Yes
RMNH.MOL.342498	Lbul_AMT22_53_05	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952850	MN950638	Yes
RMNH.MOL.342499	Lbul_AMT22_53_07	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952851	MN950639	Yes
RMNH.MOL.342500	Lbul_AMT22_53_08	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952852	MN950640	Yes
RMNH.MOL.342501	Lbul_AMT22_53_09	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952853	MN950641	
RMNH.MOL.342502	Lbul_AMT22_53_10	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952854	MN950642	Yes
RMNH.MOL.342503	Lbul_AMT22_53_11	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012			Yes
RMNH.MOL.342504	Lbul_AMT22_53_12	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012			Yes
RMNH.MOL.342505	Lbul_AMT22_53_13	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952855	MN950643	
RMNH.MOL.342506	Lbul_AMT22_53_14	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952856	MN950644	Yes
RMNH.MOL.342507	Lbul_AMT22_53_15	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952857	MN950645	Yes
RMNH.MOL.342508	Lbul_AMT22_53_16	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952858	MN950646	
RMNH.MOL.342509	Lbul_AMT22_53_17	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952859	MN950647	Yes
RMNH.MOL.342510	Lbul_AMT22_53_18	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012			Yes
RMNH.MOL.342511	Lbul_AMT22_53_20	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952860	MN950648	Yes
RMNH.MOL.342512	Lbul_AMT22_53_21	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952861	MN950649	Yes
RMNH.MOL.342513	Lbul_AMT22_53_22	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952862	MN950650	
RMNH.MOL.342514	Lbul_AMT22_53_39	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952863	MN950651	
RMNH.MOL.342515	Lbul_AMT22_53_40	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952864	MN950652	Yes
RMNH.MOL.342516	Lbul_AMT22_53_41	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952865	MN950653	
RMNH.MOL.342517	Lbul_AMT22_53_42	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952866	MN950654	
RMNH.MOL.342518	Lbul_AMT22_53_43	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952867	MN950655	
RMNH.MOL.342519	Lbul_AMT22_53_44	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952868	MN950656	
RMNH.MOL.342520	Lbul_AMT22_53_45	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952869	MN950657	
RMNH.MOL.342521	Lbul_AMT22_53_46	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952870	MN950658	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342522	Lbul_AMT22_53_48	Atlantic	AMT22_53	-20.10	-24.52	6/11/2012	MN952871	MN950659	
RMNH.MOL.342523	Lbul_AMT22_55_01	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952872	MN950660	Yes
RMNH.MOL.342524	Lbul_AMT22_55_02	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952873	MN950661	
RMNH.MOL.342525	Lbul_AMT22_55_03	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952874	MN950662	
RMNH.MOL.342526	Lbul_AMT22_55_04	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952875	MN950663	
RMNH.MOL.342527	Lbul_AMT22_55_05	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952876	MN950664	
RMNH.MOL.342528	Lbul_AMT22_55_06	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952877	MN950665	Yes
RMNH.MOL.342529	Lbul_AMT22_55_07	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952878	MN950666	
RMNH.MOL.342530	Lbul_AMT22_55_08	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952879	MN950667	
RMNH.MOL.342531	Lbul_AMT22_55_10	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952880	MN950668	Yes
RMNH.MOL.342532	Lbul_AMT22_55_12	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952881	MN950669	
RMNH.MOL.342533	Lbul_AMT22_55_13	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952882	MN950670	
RMNH.MOL.340292	Lbul_AMT22_55_14	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MK642914	MN950671	
RMNH.MOL.342535	Lbul_AMT22_55_15	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952883	MN950672	
RMNH.MOL.342536	Lbul_AMT22_55_16	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952884	MN950673	
RMNH.MOL.342537	Lbul_AMT22_55_17	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952885	MN950674	
RMNH.MOL.342538	Lbul_AMT22_55_18	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952886	MN950675	
RMNH.MOL.342539	Lbul_AMT22_55_19	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952887	MN950676	
RMNH.MOL.342540	Lbul_AMT22_55_20	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952888	MN950677	
RMNH.MOL.342541	Lbul_AMT22_55_21	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952889	MN950678	
RMNH.MOL.342542	Lbul_AMT22_55_22	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952890	MN950679	
RMNH.MOL.342543	Lbul_AMT22_55_23	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952891	MN950680	
RMNH.MOL.342544	Lbul_AMT22_55_24	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952892	MN950681	
RMNH.MOL.342545	Lbul_AMT22_55_25	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952893	MN950682	Yes
RMNH.MOL.342546	Lbul_AMT22_55_26	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952894	MN950683	Yes
RMNH.MOL.342547	Lbul_AMT22_55_27	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952895	MN950684	
RMNH.MOL.342548	Lbul_AMT22_55_28	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952896	MN950756	
RMNH.MOL.342549	Lbul_AMT22_55_29	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952897	MN950757	Yes
RMNH.MOL.342550	Lbul_AMT22_55_30	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952898	MN950685	
RMNH.MOL.342551	Lbul_AMT22_55_31	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952899	MN950758	
RMNH.MOL.342552	Lbul_AMT22_55_32	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952900	MN950759	
RMNH.MOL.342553	Lbul_AMT22_55_33	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952901	MN950760	
RMNH.MOL.342554	Lbul_AMT22_55_34	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952902	MN950686	
RMNH.MOL.342555	Lbul_AMT22_55_36	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952903	MN950687	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342556	Lbul_AMT22_55_40	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952904	MN950688	
RMNH.MOL.342557	Lbul_AMT22_55_41	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952905	MN950689	
RMNH.MOL.342558	Lbul_AMT22_55_42	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952906	MN950690	
RMNH.MOL.342559	Lbul_AMT22_55_43	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952907	MN950691	
RMNH.MOL.342560	Lbul_AMT22_55_44	Atlantic	AMT22_55	-22.95	-25.00	8/11/2012	MN952908	MN950692	
RMNH.MOL.342561	Lbul_AMT22_60_01	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952909	MN950693	
RMNH.MOL.342562	Lbul_AMT22_60_02	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952910	MN950694	
RMNH.MOL.342563	Lbul_AMT22_60_03	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012		MN950761	
RMNH.MOL.342564	Lbul_AMT22_60_04	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952911	MN950695	
RMNH.MOL.342565	Lbul_AMT22_60_05	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952912	MN950696	
RMNH.MOL.342566	Lbul_AMT22_60_08	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952913	MN950697	
RMNH.MOL.342567	Lbul_AMT22_60_09	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952914	MN950698	
RMNH.MOL.342568	Lbul_AMT22_60_10	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952915	MN950699	
RMNH.MOL.342569	Lbul_AMT22_60_11	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952916	MN950700	
RMNH.MOL.342570	Lbul_AMT22_60_12	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952917	MN950701	
RMNH.MOL.342571	Lbul_AMT22_60_14	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952918	MN950702	
RMNH.MOL.342572	Lbul_AMT22_60_15	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952919	MN950703	
RMNH.MOL.342573	Lbul_AMT22_60_16	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952920	MN950704	
RMNH.MOL.342574	Lbul_AMT22_60_17	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952921	MN950705	
RMNH.MOL.342575	Lbul_AMT22_60_18	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952922	MN950706	
RMNH.MOL.342576	Lbul_AMT22_60_19	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952923	MN950707	
RMNH.MOL.342577	Lbul_AMT22_60_20	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952924	MN950708	
RMNH.MOL.342578	Lbul_AMT22_60_21	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952925	MN950709	
RMNH.MOL.342579	Lbul_AMT22_60_24	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952926	MN950710	
RMNH.MOL.342580	Lbul_AMT22_60_25	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012			Yes
RMNH.MOL.342581	Lbul_AMT22_60_26	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012			Yes
RMNH.MOL.342582	Lbul_AMT22_60_27	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012			Yes
RMNH.MOL.342583	Lbul_AMT22_60_28	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012			Yes
RMNH.MOL.342584	Lbul_AMT22_60_29	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012			Yes
RMNH.MOL.342585	Lbul_AMT22_60_30	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952927	MN950762	Yes
RMNH.MOL.342586	Lbul_AMT22_60_31	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012		MN950763	Yes
RMNH.MOL.342587	Lbul_AMT22_60_32	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012		MN950764	
RMNH.MOL.342588	Lbul_AMT22_60_33	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012		MN950765	Yes
RMNH.MOL.342589	Lbul_AMT22_60_34	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012		MN950766	

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342590	Lbul_AMT22_60_36	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952928	MN950767	
RMNH.MOL.342591	Lbul_AMT22_60_37	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952929	MN950768	
RMNH.MOL.342592	Lbul_AMT22_60_38	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952930	MN950769	
RMNH.MOL.342593	Lbul_AMT22_60_40	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952931	MN950711	
RMNH.MOL.342594	Lbul_AMT22_60_41	Atlantic	AMT22_60	-30.17	-27.90	12/11/2012	MN952932		
RMNH.MOL.342595	Lbul_AMT22_62_03	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952933	MN950712	
RMNH.MOL.342596	Lbul_AMT22_62_04	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952934	MN950713	
RMNH.MOL.342597	Lbul_AMT22_62_08	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952935	MN950714	
RMNH.MOL.342598	Lbul_AMT22_62_12	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952936	MN950770	
RMNH.MOL.342599	Lbul_AMT22_62_13	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952937		
RMNH.MOL.342600	Lbul_AMT22_62_15	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952938	MN950715	
RMNH.MOL.342601	Lbul_AMT22_62_16	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952939	MN950716	
RMNH.MOL.342602	Lbul_AMT22_62_25	Atlantic	AMT22_62	-34.12	-33.50	14/11/2012	MN952940	MN950717	
RMNH.MOL.342603	Lbul_AMT22_64A_03	Atlantic	AMT22_64A	-35.87	-36.00	14/11/2012	MN952941	MN950718	
RMNH.MOL.342604	Lbul_AMT22_64A_07	Atlantic	AMT22_64A	-35.87	-36.00	14/11/2012	MN952942		
RMNH.MOL.342605	Lbul_AMT22_64A_09	Atlantic	AMT22_64A	-35.87	-36.00	14/11/2012	MN952943	MN950719	
RMNH.MOL.342606	Lbul_KH1110_08_01	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952944	MN950771	Yes
RMNH.MOL.342607	Lbul_KH1110_08_02	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952945		Yes
RMNH.MOL.342608	Lbul_KH1110_08_03	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952946	MN950772	Yes
RMNH.MOL.342609	Lbul_KH1110_08_04	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952947	MN950720	Yes
RMNH.MOL.342610	Lbul_KH1110_08_05	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952948	MN950773	Yes
RMNH.MOL.342611	Lbul_KH1110_08_06	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952949	MN950774	Yes
RMNH.MOL.342612	Lbul_KH1110_08_07	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952950		Yes
RMNH.MOL.342613	Lbul_KH1110_08_08	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952951	MN950775	Yes
RMNH.MOL.342614	Lbul_KH1110_08_09	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952952	MN950721	Yes
RMNH.MOL.342615	Lbul_KH1110_08_10	Pacific	KH1110_08	22.78	-158.10	19/12/2011			Yes
RMNH.MOL.342616	Lbul_KH1110_08_11	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952953	MN950722	Yes
RMNH.MOL.342617	Lbul_KH1110_08_12	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952954	MN950723	Yes
RMNH.MOL.342618	Lbul_KH1110_08_13	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952955	MN950724	Yes
RMNH.MOL.342619	Lbul_KH1110_08_14	Pacific	KH1110_08	22.78	-158.10	19/12/2011			Yes
RMNH.MOL.342620	Lbul_KH1110_08_15	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952956	MN950725	Yes
RMNH.MOL.342621	Lbul_KH1110_08_16	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952957	MN950726	Yes
RMNH.MOL.342622	Lbul_KH1110_08_17	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952958	MN950727	Yes
RMNH.MOL.342623	Lbul_KH1110_08_18	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952959	MN950728	Yes

TABLE S1 Continued.

Specimen voucher	Specimen ID	Ocean	Station	Latitude	Longitude	Collection date	COI	28S	Included in morphometrics
RMNH.MOL.342624	Lbul_KH1110_08_19	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952960	MN950729	Yes
RMNH.MOL.342625	Lbul_KH1110_08_20	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952961	MN950730	Yes
RMNH.MOL.342626	Lbul_KH1110_08_21	Pacific	KH1110_08	22.78	-158.10	19/12/2011			Yes
RMNH.MOL.342627	Lbul_KH1110_08_22	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952962	MN950731	Yes
RMNH.MOL.342628	Lbul_KH1110_08_23	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952963	MN950732	Yes
RMNH.MOL.342629	Lbul_KH1110_08_24	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952964	MN950733	Yes
RMNH.MOL.342630	Lbul_KH1110_08_25	Pacific	KH1110_08	22.78	-158.10	19/12/2011	MN952965	MN950734	Yes

Table S2 Mantel tests for pairwise Φ_{ST} and geographic distance across the entire Atlantic transect as well as different population groups. Significant Φ_{ST} values after Bonferroni correction ($\alpha = 0.05$, $p < 0.005$) are in bold.

Group tested	COI		28S	
	r	p	r	p
Atlantic	0.52	**0.0012**	-0.0942	0.77
North + Equatorial	0.117	0.269	0.759	**0.003**
South	0.0968	0.418	0.732	0.0494
North	-	-	0.359	0.202
Equatorial + South	-	-	0.428	0.0317
North + South	-	-	0.339	0.0301
Equatorial	-	-	-0.814	0.832

Table S3 Details of 13 expatriate individuals of *Limacina bulimoides* based on mitochondrial cytochrome *c* oxidase I (COI) haplogroup (see Figures 2 and 3), including information of nuclear 28S haplotype and life stage (n.a. is not available). See also Figure S4.

Expatriate	Site	COI haplogroup	28S haplotype	Life stage
Lbul_AMT22_19_44	19 (North)	2	most common	juvenile
Lbul_AMT22_25_07	25 (North)	2	one step away from most common	juvenile
Lbul_AMT22_49_01	49 (Eq.)	2	most common	adult
Lbul_AMT22_49_05	49 (Eq.)	2	most common	n.a.
Lbul_AMT22_49_06	49 (Eq.)	2	most common	n.a.
Lbul_AMT22_49_36	49 (Eq.)	2	most common	juvenile
Lbul_AMT22_51_14	51 (South)	1	most common	juvenile
Lbul_AMT22_53_01	53 (South)	1	most common	adult
Lbul_AMT22_53_17	53 (South)	1	most common	adult
Lbul_AMT22_55_08	55 (South)	1	most common	n.a.
Lbul_AMT22_55_36	55 (South)	1	most common	juvenile
Lbul_AMT22_55_41	55 (South)	1	most common	juvenile
Lbul_AMT22_64A_07	64 (South)	1	n.a.	n.a.

Table S4 Repeatability analysis (N = 30 *Limacina bulimoides* individuals x two images each) by comparison of centroid size and relative warps (RW) through intraclass correlation coefficient (ICC). Repeatable parameters, defined as having an ICC of more than 0.75, are labelled in bold.

Parameter	ICC	95% CI	% explained
Centroid	0.9983	0.9962 - 0.9992	-
RW1	0.9948	0.9891 - 0.9975	81.99
RW2	0.8581	0.7247 - 0.9297	5.3
RW3	0.6626	0.3994 - 0.8243	4.17
RW4	0.8537	0.7126 - 0.928	3.22
RW5	0.6917	0.4435 - 0.8408	2.26
RW6	0.8276	0.6715 - 0.9138	0.86
RW7	0.7906	0.6075 - 0.8942	0.71

TABLE S5 Results of ANOVA and Tukey HSD multiple pairwise comparisons of Canonical Variate 1 (see FIGURE S2) across the three Atlantic population groups of *Limacina bulimoides* (North, Equatorial and South). Significant comparisons are indicated in bold.

Group	n	Mean	SD	Tukey HSD comparisons		
				North	Equatorial	South
North	70	-3.19102	0.913883	-	-	-
Equatorial	9	-3.35496	1.255741	0.88871	-	-
South	32	-1.59519	1.103622	**0.00000**	**0.00003**	-

3

Novel genomic resources for shelled pteropods: a draft genome and target capture probes for *Limacina bulimoides*, tested for cross-species relevance

L.Q. Choo*, T.M.P. Bal*, M. Choquet, I. Smolina, P. Ramos-Silva, F. Marlétaz, M. Kopp, G. Hoarau, K.T.C.A. Peijnenburg

*Shared first authorship

Abstract

Background: Pteropods are planktonic gastropods that are considered as bio-indicators to monitor impacts of ocean acidification on marine ecosystems. In order to gain insight into their adaptive potential to future environmental changes, it is critical to use adequate molecular tools to delimit species and population boundaries and to assess their genetic connectivity. We developed a set of target capture probes to investigate genetic variation across their large-sized genome using a population genomics approach. Target capture is less limited by DNA amount and quality than other genome-reduced representation protocols, and has the potential for application on closely related species based on probes designed from one species.

Results: We generated the first draft genome of a pteropod, *Limacina bulimoides*, resulting in a fragmented assembly of 2.9 Gbp. Using this assembly and a transcriptome as a reference, we designed a set of 2,899 genome-wide target capture probes for *L. bulimoides*. The set of probes includes 2,812 single copy nuclear targets, the 28S rDNA sequence, ten mitochondrial genes, 35 candidate biomineralisation genes, and 41 non-coding regions. The capture reaction performed with these probes was highly efficient with 97% of the targets recovered on the focal species. A total of 137,938 single nucleotide polymorphism markers were obtained from the captured sequences across a test panel of nine individuals. The probes set was also tested on four related species: *L. trochiformis*, *L. lesueurii*, *L. helicina*, and *Heliconoides inflatus*, showing an exponential decrease in capture efficiency with increased genetic distance from the focal species. Sixty-two targets were sufficiently conserved to be recovered consistently across all five species.

Conclusion: The target capture protocol used in this study was effective in capturing genome-wide variation in the focal species *L. bulimoides*, suitable for population genomic analyses, while providing insights into conserved genomic regions in related species. The present study provides new genomic resources for pteropods and supports the use of target capture-based protocols to efficiently characterise genomic variation in small non-model organisms with large genomes.

Keywords

targeted sequencing, exon capture, genome, non-model organism, marine zooplankton

This chapter is published as

Choo, L.Q., Bal, T.M.P., Choquet, M., Smolina, I., Ramos-Silva, P., Marlétaz, F., Kopp, M., Hoarau, G., Peijnenburg, K.T.C.A., 2020. Novel genomic resources for shelled pteropods: A draft genome and target capture probes for *Limacina bulimoides*, tested for cross-species relevance. BMC Genomics 21, 11.

Background

Shelled pteropods are marine, holoplanktonic gastropods commonly known as 'sea butterflies', with body size ranging from a few millimetres (most species) to 1-2 centimetres (Lalli and Gilmer, 1989). They constitute an important part of the global marine zooplankton assemblage (e.g., Bednaršek et al., 2012a; Burridge et al., 2017a) and are a dominant component of the zooplankton biomass in polar regions (Hunt et al., 2008; Manno et al., 2017). Pteropods are also a key functional group in marine biogeochemical models because of their high abundance and dual role as planktonic consumers as well as calcifiers (e.g., Bé and Gilmer, 1977; Buitenhuis et al., 2019). Shelled pteropods are highly sensitive to dissolution under decreasing oceanic pH levels (Bednaršek et al., 2012a; Comeau et al., 2012; Lischka et al., 2011) because their shells are made of aragonite, an easily soluble form of calcium carbonate (Mucci, 1983). Hence, shelled pteropods may be the 'canaries in an oceanic coal mine', signalling the early effects of ocean acidification on marine organisms caused by anthropogenic releases of CO_2 (Bednaršek et al., 2017b; Manno et al., 2017). In spite of their vulnerability to ocean acidification and their important trophic and biogeochemical roles in the global marine ecosystem, little is known about their resilience towards changing conditions (Manno et al., 2017).

Given the large population sizes of marine zooplankton in general, including shelled pteropods, adaptive responses to even weak selective forces may be expected as the loss of variation due to genetic drift should be negligible (Peijnenburg and Goetze, 2013). Furthermore, the geographic scale over which gene flow occurs, between populations facing different environmental conditions, may influence their evolutionary potential (Sanford and Kelly, 2011) and consequently needs to be accounted for. It is thus crucial to use adequate molecular tools to delimit species and population boundaries in shelled pteropods.

So far, genetic connectivity studies in shelled pteropods have been limited to the use of single molecular markers. Analyses using the mitochondrial cytochrome oxidase subunit I (COI) and the nuclear 28S genes have revealed dispersal barriers at basin-wide scales in pteropod species belonging to the genera *Cuvierina* and *Diacavolinia* (Burridge et al., 2015, 2019). For *Limacina helicina*, the Arctic and Antarctic populations were discovered to be separate species through differences in the COI gene (Hunt et al., 2010; Sromek et al., 2015). However, the use of a few molecular markers has often been insufficient to detect subtle patterns of population structure expected in high gene flow species such as marine fish and zooplankton (Bucklin et al., 2018; Gaggiotti et al., 2009; Waples, 1998). In order to identify potential barriers to dispersal, we need to sample a large number of loci across the genome, which is possible due to recent developments in next-generation sequencing (NGS) technologies (De Wit et al., 2015; McCormack et al., 2013).

Here, we chose a genome reduced-representation method to characterise genome-wide variation in pteropods because of their potentially large genome sizes and small amount of input DNA per individual. In species with large genomes,

as reported for several zooplankton groups (Bucklin et al., 2018), whole genome sequencing may not be feasible for population-level studies. Reduced-representation methods can overcome the difficulty of sequencing numerous large genomes. Two common approaches are RADseq and target capture enrichment. RADseq (Baird et al., 2008), which involves the enzymatic fragmentation of genomic DNA followed by the selective sequencing of the regions flanking the restriction sites of the used enzyme(s), is attractive for non-model organisms as no prior knowledge of the genome is required. However, RADseq protocols require between 50 ng and 1 µg of high-quality DNA, with higher amounts being recommended for better performance (Andrews et al., 2016), and has faced substantial challenges in other planktonic organisms (e.g., Choquet et al., 2019; Deagle et al., 2015). Furthermore, RADseq may not be cost efficient for species with large genomes (Choquet et al., 2019). Target capture enrichment (Glenn and Faircloth, 2016; Jones and Good, 2016; Mamanova et al., 2010) overcomes this limitation in DNA starting amount and quality, by using single-stranded DNA probes to selectively hybridise to specific genomic regions that are then recovered and sequenced (Gnirke et al., 2009). It has been successfully tested on large genomes with just 10 ng of input DNA (Chung et al., 2016) as well as degraded DNA from museum specimens (Bi et al., 2013; Blaimer et al., 2016; Kollias et al., 2015; McCormack et al., 2016). Additionally, the high sequencing coverage of targeted regions allows rare alleles to be detected (Chung et al., 2016).

Prior knowledge of the genome is required for probe design, however, this information is usually limited for non-model organisms. Currently, there is no pteropod genome available that can be used for the design of genome-wide target capture probes. The closest genome available is from the sister group of pteropods, Anaspidea (*Aplysia californica*, NCBI reference: PRJNA13635; Broad Institute, 2009), but it is too distant to be a reference, as pteropods have diverged from other gastropods since at least the Late Cretaceous (Burridge et al., 2017b).

In this study, we designed target capture probes for the shelled pteropod *Limacina bulimoides* based on the method developed in Choquet et al. (2019), to address population genomic questions using a genome-wide approach. We obtained the draft genome of *L. bulimoides* to develop a set of target capture probes, and tested the success of these probes through the number of single nucleotide polymorphisms (SNPs) recovered in the focal species. *L. bulimoides* was chosen as the probe-design species because it is an abundant species with a worldwide distribution across environmental gradients in subtropical and tropical oceans. The probes were also tested on four related species within the Limacinoidea superfamily (coiled-shell pteropods) to assess their cross-species effectiveness. Limacinoid pteropods have a high abundance and biomass in the world's oceans (Bé and Gilmer, 1977; Bednaršek et al., 2012a; Burridge et al., 2017b) and have been the focus of most ocean acidification research to date (e.g., Bednaršek et al., 2012a; Maas et al., 2018; Moya et al., 2016).

Table 1 Summary of draft genome statistics for *Limacina bulimoides*.

Assembly statistics	Value
Estimated total genome size	4,801,432,559 bp
Total assembly size	2,901,932,435 bp
Number of scaffolds	
>= 0 bp	3,735,734
>= 1000 bp	802,059
>= 5000 bp	3,890
>= 10,000 bp	116
>= 25,000 bp	6
>= 50,000 bp	3
N50	893 bp
L50	994,289
Smallest scaffold	200 bp
Largest scaffold	197,255 bp
Percentage of N's	0.3307
GC content, %	34.08

Results

Draft genome assembly

We obtained a draft genome of *L. bulimoides* (NCBI: SWLX00000000) from 108 Gb of Illumina data sequenced as 357 million pairs of 150 base pair (bp) reads. As a first pass in assessing genomic data completeness, a k-mer spectrum analysis was done with JELLYFISH version 1.1.11 (Marçais and Kingsford, 2011). It did not show a clear coverage peak, making it difficult to estimate total genome size with the available sequencing data (Appendix S1). Because distinguishing sequencing error from a coverage peak is difficult below 10-15x coverage, it is likely that the genome coverage is below 10-15x, suggesting a genome size of at least 6-7 Gb. The reads were assembled using the *de novo* assembler MaSuRCA (Zimin et al., 2013) into 3.86 million contigs with a total assembly size of 2.9 Gbp (N50 = 851 bp, L50 = 1,059,429 contigs). The contigs were further assembled into 3.7 million scaffolds with a GC content of 34.08% (Table 1). Scaffolding resulted in a slight improvement, with an increase in the N50 to 893 bp and a decrease in the L50 to 994,289 contigs. Based on the hash of error corrected reads in MaSuRCA, the total haploid genome size was estimated at 4,801,432,459 bp (4.8 Gbp). Therefore, a predicted 60.4% of the complete genome was sequenced.

Genome completeness based on the assembled draft genome was measured in BUSCO version 3.0.1 (Simão et al., 2015) and resulted in the detection of 60.2% of near universal orthologues that were either completely or partially present in the draft genome of *L. bulimoides* (Table 2). This suggests that around 40% of gene information is missing or may be too divergent from the BUSCO sets (Simão et al., 2015). Although the use of BUSCO on a fragmented genome may not give reliable estimates as orthologues may be partially represented within scaffolds that are too short for a positive gene prediction, this percentage of near-universal orthologues coincides with the estimate of genome size by MaSuRCA.

Table 2 Summary of BUSCO analysis showing the number of metazoan near universal orthologues that could be detected in the draft genome of *Limacina bulimoides*.

	Present in draft genome
Complete	296 (30.3%)
Complete and single-copy	262 (26.8%)
Complete and duplicated	34 (3.5%)
Fragmented	292 (29.9%)
Missing	390 (39.8%)
Total BUSCO groups searched	978

We also compared the draft genome to a previously generated transcriptome of *L. bulimoides* (NCBI: SRR10527256) (Peijnenburg et al., 2020) to assess the completeness of the coding sequences and aid in the design of capture probes. The transcriptome consisted of 116,995 transcripts, with an N50 of 555 bp. Even though only ~60% of the genome was assembled, 79.8% (93,306) of the transcripts could be mapped onto it using the splice-aware mapper GMAP version 2017-05-03 (Wu and Watanabe, 2005). About half of the transcripts (46,701 transcripts) had single mapping paths and the other half (46,605 transcripts) had multiple mapping paths. These multiple mapping paths are most likely due to the fragmentation of genes over at least two different scaffolds, but may also indicate multi-copy genes or transcripts with multiple spliced isoforms. Of the singly mapped transcripts, 8,374 mapped to a scaffold that contained two or more distinct exons separated by introns. Across all the mapped transcripts, 73,719 were highly reliable with an identity score of 95% or higher.

Target capture probes design and efficiency

A set of 2,899 genome-wide probes, ranging from 105 to 1,095 bp, was designed for *L. bulimoides*. This includes 2,812 single copy nuclear targets of which 643 targets were previously identifed as conserved pteropod orthologs (Peijnenburg et al., 2020), the 28S rDNA sequence, 10 known mitochondrial genes, 35 candidate biomineralisation genes (Mann and Jackson, 2014; Ramos-Silva and Marin, 2016), and 41 randomly selected non-coding regions (see Methods). The set of probes worked very well on the focal species *L. bulimoides*. Of the targeted regions, 97% (2,822 of 2,899 targets) were recovered across a test panel of nine individuals (Table 3) with 137,938 SNPs (Table 4) identified across these targeted regions. Each SNP was present in at least 80% of *L. bulimoides* individuals (also referred to as genotyping rate) with a minimum read depth of 5x. Coverage was sufficiently high for SNP calling (Figure 3) and 87% of the recovered targets (2,446 of the 2,822 targets) had a sequence depth of 15x or more across at least 90% of their bases (Figure 1A). Of the 2,822 targets, 643 targets accounted for 50% of the total aligned reads in *L. bulimoides* (Figure S2A in Appendix S2). For *L. bulimoides*, SNPs were found in all categories of targets, including candidate biomineralisation genes, non-coding regions, conserved pteropod orthologues, nuclear 28S and other coding

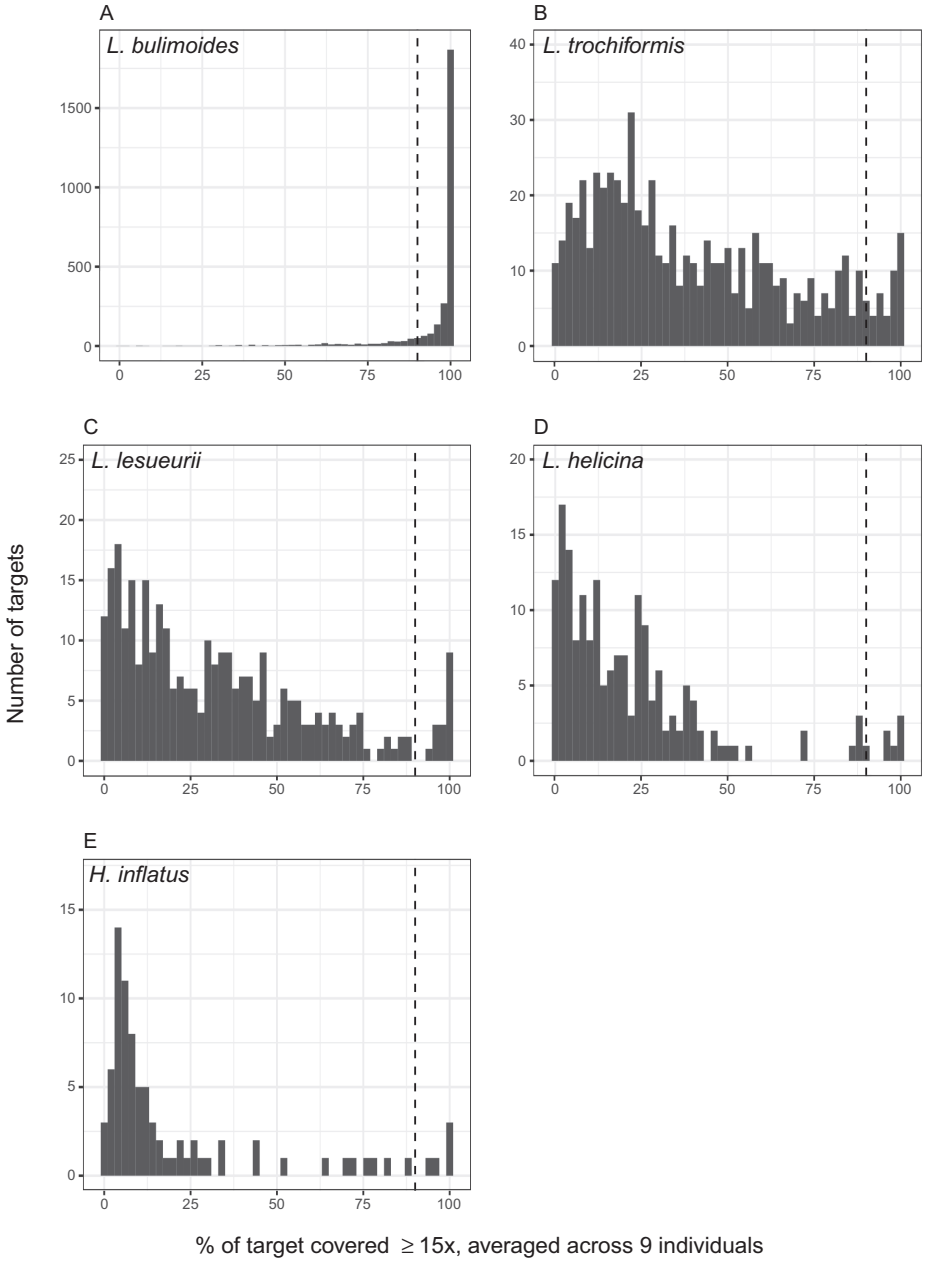


FIGURE 1 Number of recovered targets plotted against average proportion of bases in each target, with at least 15x sequencing coverage averaged across nine individuals, for each for the five shelled pteropod species (*Limacina bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. helicina*, and *Heliconoides inflatus*). Bars on the right of the dashed vertical line represent the number of targets where more than 90% of the bases in each target was sequenced with ≥15x depth. Note the differences in y-axes between the plots. There is no peak at one SNP for *L. bulimoides* (APPENDIX S5).

TABLE 3 Target capture efficiency statistics, averaged ± standard deviation across nine individuals, for each of five pteropod species, including raw reads, final mapped reads, % High Quality reads (reads mapping uniquely to the targets with proper pairs), % targets covered (percentage of bases across all targets covered by at least one read), average depth (sequencing depth across all targets with reads mapped).

Species	Raw reads (x1,000)	Final mapped reads (x1,000)	% HQ reads	% targets covered	Average depth
L. bulimoides	10,529±3,997	3,531±1,548	33.23±9.10	97.36±0.42	250±111
L. trochiformis	15,508±4,865	1,765±521	11.61±2.59	20.32±1.65	468±144
L. lesueurii	7,060±2,043	807±196	11.93±2.77	13.28±1.96	431±76.9
L. helicina	10,346±6,260	337±180	3.47±0.56	12.57±2.71	63.7±26.7
H. inflatus	3,089±1,126	66±30	2.07±0.30	8.21±3.34	31.9±14.9

TABLE 4 Number of single nucleotide polymorphism (SNPs) recovered after various filtering stages for five species of shelled pteropods. Hard-filtering was implemented in GATK3.8 VariantFiltration using the following settings: QualByDepth <2.0, FisherStrand >60.0, RMSMappingQuality <5.0, MQRankSumTest <-5.0 and ReadPositionRankSum <-5.0. The hard-filtered SNPs were subsequently filtered to keep those with a minimum site coverage of 5x and present in at least 80% of the individuals. Other filtering options were less stringent, such as a minimum depth of 2x and site presence in at least 50% of individuals.

	Hard-filtering	80% individuals, 5x depth	80% individuals, 2x depth	50% individuals, 5x depth
L. bulimoides	154,864	137,938	137,953	147,763
L. trochiformis	44,014	11,948	12,165	20,518
L. lesueurii	23,379	5,359	5,847	8,487
L. helicina	18,298	2,432	2,771	4,613
H. inflatus	13,041	1,371	1,559	2,092

TABLE 5 Number of targets with at least one single nucleotide polymorphism (based on 80% genotyping rate, 5x depth) was calculated according to category: candidate biomineralisation genes (Biomin.), conserved pteropod orthologues, mitochondrial (Mt genes), nuclear 28S, and other coding and non-coding regions for each of five pteropod species. Numbers in brackets represent the total number of targets in that category on the set of target probes designed for *Limacina bulimoides*.

Species	Biomin. (35)	Orthologues (643)	Mt genes (10)	28S (1)	Coding (2,169)	Non-coding (41)	Total (2,899)
L. bulimoides	32	635	1	1	2,140	13	2,822
L. trochiformis	7	169	3	1	436	4	620
L. lesueurii	0	90	2	1	209	0	302
L. helicina	0	52	3	1	121	0	177
H. inflatus	0	20	1	1	61	0	83

sequences (TABLE 5). Of the 10 mitochondrial genes included in the capture, surprisingly, only the COI target was recovered.

The hybridisation of the probes and targeted re-sequencing worked much less efficiently on the four related species. The percentage of targets covered by sequenced reads ranged from 8.21% (83 out of 2,899 targets) in *H. inflatus* to

20.32% (620 out of 2,899 targets) in *L. trochiformis* (Table 3). Of these, only five (*H. inflatus*) to 42 (*L. trochiformis*) targets were covered with a minimum of 15x depth across 90% of the bases (Table S1). The number of targets that accounted for 50% of the total aligned reads varied across species, with 4 of 620 targets for *L. trochiformis* that accounted for 50% of reads, 2 of 302 targets for *L. lesueurii*, 14 of 177 targets for *L. helicina* and 5 of 83 targets for *H. inflatus* (Figure S2B-E in Appendix S2). In these four species, targeted regions corresponding to the nuclear 28S gene, conserved pteropod orthologues, mitochondrial genes and other coding sequences were obtained (Table 4). The number of mitochondrial targets recovered ranged between one and three: ATP6, COB, 16S were obtained for *L. trochiformis*, ATP6, COI for *L. lesueurii*, ATP6, COII, 16S for *L. helicina*, and only 16S for *H. inflatus*. Additionally, for *L. trochiformis*, seven biomineralisation candidates and four non-coding targeted regions were recovered. The number of SNPs ranged between 1,371 (*H. inflatus*) and 12,165 SNPs (*L. trochiformis*) based on a gentoyping rate of 80% and a minimum read depth 5x (Table 5). The maximum depth for SNPs ranged from ~150x in *H. inflatus*, *L. helicina* and *L. lesueurii* to ~375x in *L. trochiformis* (Figure 3). With less stringent filtering, such as a 50% genotyping rate, the total number of SNPs obtained per species could be increased (Table 5).

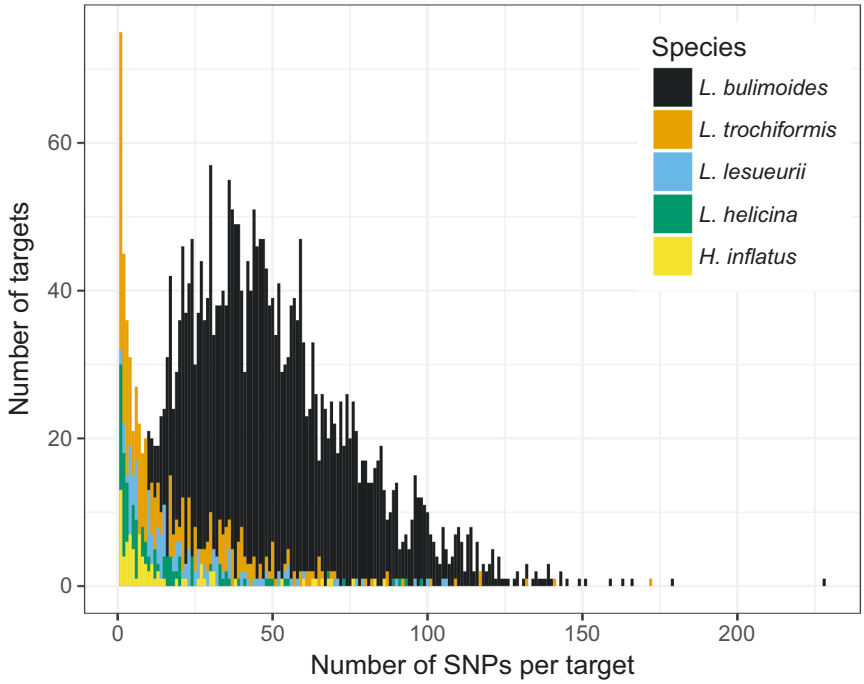


Figure 2 Number of single nucleotide polymorphisms (SNPs) per recovered target for the five pteropod species of the superfamily Limacinoidea (see legend), based on filtering settings of minimum presence in 80% of individuals with at least 5x read depth.

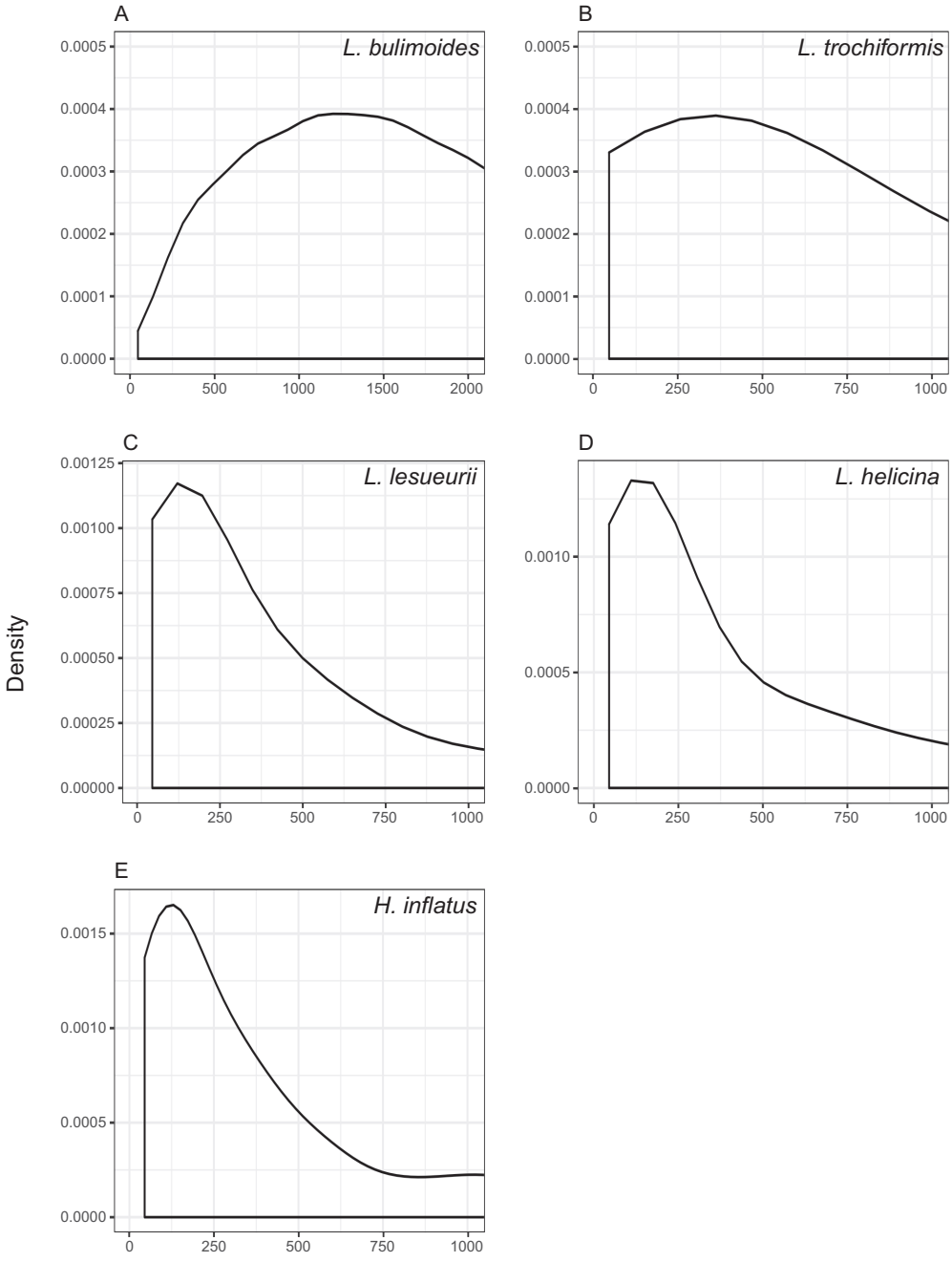


FIGURE 3 Density of single nucleotide polymorphisms (SNPs, present in 80% of individuals) plotted against coverage. The plots were truncated at coverage = 2000x for *L. bulimoides* and coverage = 1000x for the other four species. Note that minimum coverage is 45x due to filtering settings of a minimum 5x depth for nine individuals.

Across the five species of Limacinoidea, we found an exponential decrease in the efficiency of the targeted re-sequencing congruent with the genetic distance from the focal species *L. bulimoides*. Only 62 targets were found in common across all five species, comprising 14 conserved pteropod orthologues, 47 coding regions, and a 700 bp portion of the 28S nuclear gene. Based on the differences in profiles of number of SNPs per target and total number of SNPs, the hybridisation worked differently between the focal and non-focal species. In *L. bulimoides*, the median number of SNPs per target was 45, whereas in the remaining four species, most of the targets had only one SNP and the median number of SNPs per target was much lower: 11 for *L. trochiformis*, 10 for *L. lesueurii*, six for *L. helicina*, and seven for *H. inflatus*. The number of SNPs per target varied between one and more than 200 across the targets (Figure 2). With an increase in genetic distance from *L. bulimoides,* the total number of SNPs obtained across the five shelled pteropod species decreased exponentially (Figure 4). There was an initial 10-fold decrease in number of SNPs between *L. bulimoides* and *L. trochiformis* with a maximum likelihood (ML) distance of 0.07 nucleotide substitutions per base between them. The subsequent decrease in number of SNPs was smaller in *L. lesueurii* (ML distance from *L. bulimoides*, subsequently ML dist = 0.11), *L. helicina* (ML dist = 0.18) and *H. inflatus* (ML dist = 0.29).

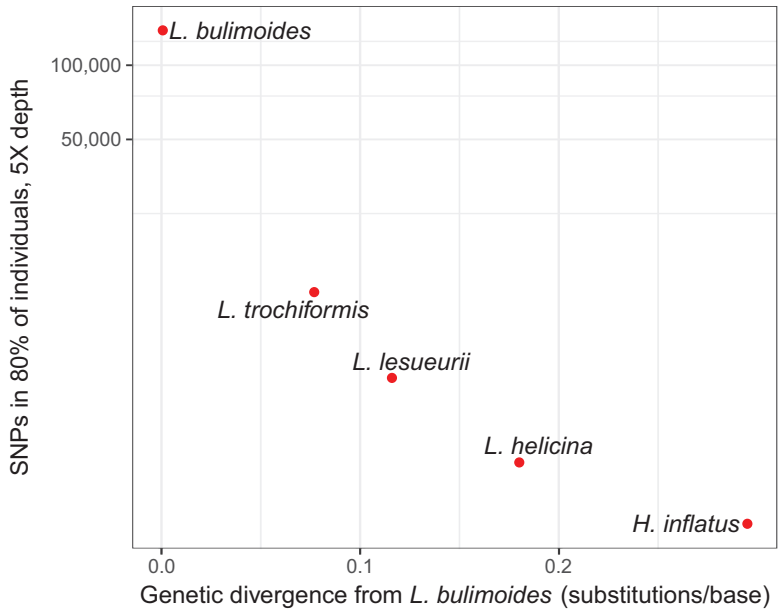


Figure 4 Log-scaled number of SNPs against genetic divergence from the focal species *Limacina bulimoides* shows that there is a sharp reduction in the SNPs recovered with genetic distance.

Discussion

First draft genome for pteropods

To assess the genetic variability and degree of population connectivity in coiled-shell pteropods, we designed a set of target capture probes based on partial genomic and transcriptomic resources. As a first step, we *de novo* assembled a draft genome for *L. bulimoides*, the first for a planktonic gastropod. We obtained an assembly size of 2.9 Gbp but the prediction of genome size together with the prediction of genome completeness suggest that only ~60% of the genome was sequenced. Therefore, we postulate that the genome size of *L. bulimoides* is indeed larger than the assembly size, and estimate it at 6-7 Gbp. In comparison, previously sequenced molluscan genomes have shown a wide variation in size across species, ranging from 412 Mbp in the giant owl limpet (*Lottia gigantea)* (Simakov et al., 2013) to 2.7 Gbp in the Californian two-spot octopus (*Octopus bimaculoides)* (Albertin et al., 2015). The closest species to pteropods which has a sequenced genome is *Aplysia californica*, with a genome size of 927 Mbp (Genbank accession assembly: GCA_000002075.2) (Broad Institute, 2009; Sayers et al., 2019). Further, when considering marine gastropod genome size estimates in the Animal Genome Size Database (Gregory, 2019), genome sizes range from 430 Mbp to 5.88 Gbp with an average size of 1.86 Gbp. Hence, it appears that *L. bulimoides* has a larger genome size than most other gastropods.

Despite moderate sequencing efforts, our genome is highly fragmented. Increasing the sequencing depth could result in some improvements, although other sequencing methods will be required to obtain a better genome. Roughly 350 million paired-end (PE) reads were used for the *de novo* assembly, but 50% of the assembly is still largely unresolved with fragments smaller than 893 bp. The absence of peaks in the k-mer distribution histogram and low mean coverage of the draft genome may indicate insufficient sequencing depth caused by a large total genome size, and/or high heterozygosity which complicates the assembly. In the 1.6 Gbp genome of another gastropod, the big-ear radix, *Radix auricularia*, approximately 70% of the content consisted of repeats (Schell et al., 2017). As far as we know, high levels of repetitiveness within molluscan genomes are common (Takeuchi, 2017), and also makes *de novo* assembly using only short reads challenging (Treangen and Salzberg, 2012). In order to overcome this challenge, genome sequencing projects should combine both short and long reads to resolve repetitive regions that span across short reads (Koren et al., 2012; Rice and Green, 2018). Single molecule real time (SMRT) sequencing techniques which produce long reads recommend substantial DNA input, although some recent developments in library preparation techniques have lowered the required amount of DNA (Kingan et al., 2018). These SMRT techniques also tend to be high in cost, which may be a limiting factor when choosing between sequencing methods. Constant new developments in sequencing-related technologies may soon bring the tools needed to achieve proper genome assembly even for small-sized organisms with

large genomes. Potential methods to improve current shotgun assemblies include 10x Genomics linked-reads (10X Genomics, 2019) that uses microfluidics to leverage barcoded subpopulations of genomic DNA or Hi-C (Belton et al., 2012), which allow sequences in close physical proximity to be identified as linkage groups and enable less fragmented assemblies.

Target capture probes for *Limacina bulimoides*

Our results show that generating a draft genome and transcriptome to serve as a reference in the design of target capture probes is a promising and cost-effective approach to allow population genomics studies in non-model species of small sizes. Despite the relatively low N50 of the assembled genome, we were able to map 79.8% of the transcript sequences onto it. The combined use of the transcriptome and fragmented genome allowed us to identify the expressed genomic regions reliably and include intronic regions, which may have contributed to the probe hybridisation success (Suren et al., 2016). In addition, the draft genome was useful in obtaining single-copy regions. This allowed us to filter out multi-copy regions at the probe design step, and hence reducing the number of non-target matches during the capture procedure.

The target capture was highly successful in the focal species *L. bulimoides*, with more than 130,000 SNPs recovered across nine individuals (Figure 3). Coverage of reads across the recovered targets was somewhat variable (Figure S2A in Appendix S2), although the SNPs were obtained from the large proportion of sufficiently well-covered targets (>15x, Table 4; Table S1) and thus, can provide reliable genomic information for downstream analyses, such as delimiting population structure. The high number of SNPs may be indicative of high levels of genetic variation, congruent with predictions for marine zooplankton with large population sizes (Peijnenburg and Goetze, 2013). The number of SNPs recovered (Table 4) and percentage of properly paired reads mapping uniquely to the targets (Table 3) are comparable to the results from a similar protocol on copepods (Choquet et al., 2019).

Targets corresponding to candidate biomineralisation genes and mitochondrial genes were less successfully recovered compared to conserved pteropod orthologues and other coding sequences (Table 4). This could be because biomineralisation-related gene families in molluscs are known to evolve rapidly, with modular proteins composed of repetitive, low complexity domains that are more likely to accumulate mutations due to unequal cross-over and replication slippage (Kocot et al., 2016; McDougall and Degnan, 2018). Surprisingly, only the COI gene was recovered out of the 10 mitochondrial genes included in the set of probes. This is despite the theoretically higher per cell copy number of mitochondrial than nuclear genomes (Bi et al., 2012) and thus a higher expected coverage for mitochondrial targets compared to nuclear targets. High levels of mitochondrial polymorphism among individuals of *L. bulimoides* could have further complicated the capture, resulting in low capture success of mitochondrial targets. Hyperdiversity in mitochondrial genes, with more than 5% nucleotide diversity in synonymous sites has

been reported for several animal clades, including gastropods (Fourdrilis et al., 2016; Thomaz et al., 1996) and chaetognaths (Marlétaz et al., 2017). Only 13 of the 41 non-coding targeted regions were recovered, which may indicate that these regions were also too divergent to be captured by the probes.

Cross-species relevance of target capture probes

The success of targeted re-sequencing of the four related pteropod species (*L. trochiformis*, *L. lesueuri*, *L. helicina* and *Heliconoides inflatus*) decreased exponentially with increasing genetic distance from the focal species *L. bulimoides*. Even within the same genus, divergence was sufficiently high to show an abrupt decrease in coverage (Figure 3). The number of targets whose reads accounted for 50% of all reads for each species was low (Figure S2B-E in Appendix S2), indicating that representation across the targets could be highly uneven. The number of SNPs recovered also decreased rapidly with genetic distance (Figure 4), leading to less informative sites across the genome that can be used in downstream analyses for these non-focal species. While direct comparisons are not possible due to differences in the probe design protocol and measurements used, we also see a decreasing trend in success of target capture applied with increasing levels of genetic divergence in other studies (e.g., Förster et al., 2018; Portik et al., 2016). Genetic divergence of 4-10% from the focal species resulted in an abrupt decline in coverage (e.g., Bi et al., 2012; Bragg et al., 2016). Another possible reason for the decrease in capture success is different genome sizes across the species. While we used the same amount of DNA per individual in a capture reaction, pooling different species of unknown genome sizes into the same capture reaction may have resulted in different genome copy numbers sequenced per species. Our results may thus be attributed to high levels of polymorphism and/or possible differences in genome size, both leading to ascertainment bias (Lachance and Tishkoff, 2013).

The targets that hybridised successfully and were sequenced across species were conserved genes with low levels of genetic variation. This probably indicates that high levels of genetic diversity and divergence from the focal species resulted in the targeted regions not being able to hybridise to the probes. Indeed, from the four non-focal pteropod species, most of the recovered targets had low diversity, containing only a single SNP (Figure 2). As a general rule, slowly evolving genomic regions are more likely to hybridise successfully to the probes (Bi et al., 2013; Paijmans et al., 2016). This may vary across targeted regions, as a mismatch tolerance of 40% between the baits and targeted region can still result in successful enrichment in specific cases (Li et al., 2013). While it is possible to design probes to be relevant across broader phylogenetic scales, by including conserved orthologues across the various target species (e.g., Quattrini et al., 2018; Teasdale et al., 2016), these probes are unlikely to be suitable to study population structure and estimate levels of gene flow in the focal species. Nonetheless, the low diversity targets that were recovered can be useful in resolving relationships at a deeper phylogenetic scale.

Conclusion

We show that using a combination of a draft genome and transcriptome is an efficient way to develop a database for capture probes design in species without prior genomic resources. These probes can be useful for analyses in closely related species, though cross-species hybridisation was limited to conserved targets and capture success decreased exponentially with increasing genetic distance from the focal species. Since the target capture approach can be successfully applied with low DNA input and even with poor quality or degraded DNA, this technique opens the door to population genomics of zooplankton, from recent as well as historical collections.

With more than 130,000 SNPs recovered in *L. bulimoides* and >10,000 SNPs in *L. trochiformis*, our set of probes is suitable for genome-wide genotyping in these two globally distributed pteropod species. The high and consistent coverage across targeted genomic regions increases the range of analyses that can be applied to these organisms, such as identifying dispersal barriers, inferring ancestry and demographic history, and detecting signatures of selection across the genome. The statistical strength from analysing many genomic loci overcomes the limitation of an incomplete sampling of the metapopulation (Maisano Delser et al., 2016) and increases the capacity to detect even subtle patterns in population structure. This is especially relevant in widespread marine zooplankton where there is likely to be cryptic diversity and undiscovered species (Bucklin et al., 2018; Peijnenburg and Goetze, 2013), which is essential information for species that are proposed as indicators of ocean change.

Materials and methods

Draft genome sequencing and assembly

A single adult *L. bulimoides* (1.27 mm total shell length) was used to generate a draft genome (NCBI: SWLX00000000). This individual was collected from the southern Atlantic subtropical gyre (25°44'S, 25°0'W) during the Atlantic Meridional Transect (AMT) cruise 22 in November 2012 (Appendix S3 and Table S3) and directly preserved in 95% ethanol at -20°C. Back in the lab, 147.2 ng of genomic DNA was extracted from the whole specimen using the E.Z.N.A. Insect DNA Kit (Omega Bio-Tek) with modifications to the manufacturer's protocol regarding reagents volumes and centrifuge times (Appendix S3). The extracted DNA was randomly fragmented via sonication on a S220 Focused-ultrasonicator (Covaris) targeting a peak length of approximately 350 bp. A genomic DNA library was prepared using the NEXTflex Rapid Pre-Capture Combo Kit (Bioo Scientific) following the manufacturer's protocol. Subsequently, the library was sequenced in two runs of NextSeq500 (Illumina) using mid-output v2 chips producing 150 bp PE reads.

The resulting forward and reverse sequencing reads were concatenated in two separate files and quality-checked using FastQC version 0.11.4 (https://www.bioin-

formatics.babraham.ac.uk/projects/fastqc/). Duplicated reads were removed using FastUniq version 0.11.5 (Xu et al., 2012). The remaining reads were then assembled by the MaSuRCA genome assembler version 3.2.1 (Zimin et al., 2013) using a k-mer length of 105 as this produced the least fragmented assembly compared to other assemblers (Platanus, SOAPdenovo2). Further contig extension and scaffolding were carried out by running SSPACE-Basic version 2 (Boetzer et al., 2011) requiring a minimum of three linkers and a minimum overlap of 12 bp to merge adjacent contigs (Boetzer et al., 2011). The total genome size was roughly estimated using MaSuRCA (as a by-product of calculating optimal assembly parameters), based on the size of the hash table containing all error corrected reads. A second estimate of the genome size was made by searching for k-mer peaks in sequencing reads using JELLYFISH version 1.1.11 (Marçais and Kingsford, 2011) with various k-mer lengths between 15 and 101. To assess the completeness of the generated draft genome, the in-built BUSCO metazoan dataset containing 978 near-universal orthologues of 65 species was used to search for key orthologous genes with BUSCO version 3.0.1 (Simão et al., 2015). BUSCO made use of AUGUSTUS version 3.3 (Stanke and Morgenstern, 2005) with the self-training mode utilised to predict gene models. Assembly quality was assessed with QUAST (Gurevich et al., 2013).

Target capture probes design

We designed the target capture probe set by using the draft genome and transcriptome as a reference, following the workflow recommended by Choquet et al. (2019). Firstly, we aimed to select only single-copy coding DNA sequences (CDS) in order to achieve a high specificity of the target capture probes and to reduce false-positive SNPs from multi-copy genes. We used the previously generated transcriptome of *L. bulimoides* (Peijnenburg et al., 2020) and mapped the transcript sequences of *L. bulimoides* against themselves using the splice-aware mapper GMAP version 2017-05-03 (Wu and Watanabe, 2005) with a k-mer length of 15 bp and no splicing allowed. Only unique transcripts with one mapping path were selected as potential target sequences. We then mapped these selected transcript sequences (with splicing allowed) directly to the contigs of the genomic assembly to identify expressed regions and their respective exon-intron boundaries. We selected only the subset of genomic sequences that mapped to unique transcripts with minimum pairwise identity scores of 90%. Using this approach, we selected 2,169 coding target sequences. Additionally, 643 transcripts that mapped to unique contigs in the draft genome were selected from a set of conserved orthologues from a phylogenomic analysis of pteropods (Peijnenburg et al., 2020) to give a set of 2,812 single copy coding nuclear targets. Of the 63 transcripts that showed homology to biomineralisation proteins (Mann and Jackson, 2014; Ramos-Silva and Marin, 2016), we included 35 of these candidate biomineralisation genes in the final probe set as they could be mapped to contigs in the draft genome (Supporting File 1).

Secondly, sequences of mitochondrial genes, 28S and non-coding targets were added to the baits design. A fragment of the COI gene (NCBI: MK642914), obtained

by sanger sequencing as in (Burridge et al., 2017b) was added. The other nine targets (COII, COIII, ATP6, ND2, ND3, ND6, CYB, 12S, 16S) were identified from the draft genome assembly as described hereafter. We identified a 9,039 bp contig from the fragmented assembly as a partially assembled mitochondrial genome using BLAST+ version 2.6.0 (Camacho et al., 2009) and comparing the mitochondrial genes of three related mollusc species (NCBI Bioprojects: PRJNA10682, PRJNA11892, PRJNA12057) to the draft genome. Gene annotation was then carried out on this contig using the MITOS webserver (Bernt et al., 2013) with the invertebrate genetic code and the parameters 'cut-off', 'fragment quality factor' and 'start/stop range' set to 30, 12 and 10, respectively. From this, we identified the seven protein-coding genes and the two rRNA genes as separate target sequences which we added to the probe design. Finally, we added the commonly-used nuclear 28S Sanger-sequenced fragment (NCBI: MK635470) and randomly chose 41 unique non-coding genomic regions. The final design comprised of 2,899 target sequences with a total size of 1,866,005 bp. Probe manufacturing was performed by Arbor Biosciences (MI, USA) using myBaits custom biotinylated probes of 82-mer with 2x tiling density.

Targeted sequencing of five pteropod species

We selected five shelled pteropod species from the genera *Limacina* and *Heliconoides* (superfamily Limacinoidea), including the focal species *L. bulimoides*, to evaluate the efficiency of the target capture probes on species of varying genetic relatedness. For each species, we aimed to test the capture efficiency across three sampling locations with three individuals per location (Table 6). Specimens from each species (*L. bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. helicina*, *H. inflatus*) were collected across various sites during the AMT22 and AMT24 cruises in the Atlantic and from two sites in the Pacific Ocean (Table 6 and Table S2). DNA was extracted from each individual separately using either E.Z.N.A. insect or mollusc kit (Omega Bio-Tek) with modifications to the protocol (Appendix S3). The DNA was then sheared by sonication, using a Covaris S220 ultrasonicator with the peak length set to 300 bp. This fragmented DNA was used to prepare individual libraries indexed using the NEXTflex Rapid Pre-Capture Combo Kit (Bioo Scientific). Libraries were subsequently pooled into equimolar concentrations for the capture reaction using the myBaits Custom Target Capture kit (Arbor Biosciences). Hybridisation was carried out using the myBaits protocol with the following modifications. 27 libraries of *L. bulimoides* were pooled together for one capture reaction, of which nine individuals were analysed in this study. The other four species were pooled in groups of 22-23 specimens per capture. We extended the hybridisation time to three days and performed the whole protocol twice using 4 µl and 1.5 µl of probe mix, respectively (Appendix S3). The captured library of the species *L. bulimoides* was sequenced on the NextSeq500 (Illumina) using a high-output v2 chip producing 150 bp PE reads. The captured libraries of the other species were sequenced together on the same NextSeq500 mid-output v2 chip.

Table 6 Collection details of specimens from five shelled pteropod species: *Limacina bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. helicina* and *Heliconoides inflatus*. Three individuals per site were included from localities in the Atlantic and Pacific Oceans. Latitude and longitude are presented in the decimal system, with positive values indicating North and East and negative values, South and West, respectively.

Species	Location	Latitude	Longitude	n	Collection date
L. bulimoides	South Atlantic	-18.32	-25.08	3	18/10/2014
L. bulimoides	South Atlantic	-24.45	-25.05	3	21/10/2014
L. bulimoides	South Atlantic	-27.77	-25.02	3	22/10/2014
L. trochiformis	South Atlantic	-14.67	-25.07	3	17/10/2014
L. trochiformis	South Atlantic	-18.32	-25.08	3	18/10/2014
L. trochiformis	North Pacific	22.65	-157.69	3	03/07/2017
L. lesueurii	North Atlantic	20.40	-38.61	3	24/10/2012
L. lesueurii	South Atlantic	-15.30	-25.07	3	05/11/2012
L. lesueurii	South Atlantic	-24.13	-25.00	3	09/11/2012
L. helicina	South Atlantic	-40.12	-30.92	3	26/10/2014
L. helicina	South Atlantic	-41.48	-33.87	3	27/10/2014
L. helicina	North Pacific	48.36	-126.31	3	06/03/2016
H. inflatus	North Atlantic	25.48	-39.00	3	22/10/2012
H. inflatus	South Atlantic	-8.08	-25.04	3	03/11/2012
H. inflatus	South Atlantic	-38.08	-39.31	3	16/11/2012

Assessment of target capture probes efficiency

The following pipeline of bioinformatic analyses was largely adapted from Choquet et al. (2019). Raw sequencing reads were de-multiplexed and mapped using BWA version 0.7.12 (Li, 2013) with default settings to targets concatenated with the perl script concatFasta.pl (Matz, 2019). The resulting BAM files were then cleaned and sorted using SAMtools version 1.4.1 (Li et al., 2009) to retain only the reads paired and uniquely mapped in proper pairs. With Picard version 2.18.5 (Broad Institute, 2019), duplicates were marked and removed. Coverage of targeted regions was assessed with the GATK version 3.8 (McKenna et al., 2010) DepthOfCoverage tool. Next, SNP calling was performed using GATK version 3.8 with GNU Parallel (Tange, 2011) following the recommended Variant Discovery pipeline (Auwera et al., 2013; Depristo et al., 2011) as a first trial for SNP calling in pteropods. Variants were called per individual using HaplotypeCaller with emitRefConfidence output, and the resulting gVCF files were combined according to their species with CombineGVCFs. The combined gVCF files for each species, with nine individuals each, were then genotyped in GenotypeGVCFs. SNPs were extracted from the raw variants with SelectVariants (-SelectType SNP). Given the lack of a calibration set of SNPs, the hard filters were first evaluated by plotting the density of annotation values and checking them against the planned filtering parameters. The SNPs were then hard-filtered with VariantFiltration using QualByDepth (QD) <2.0, FisherStrand (FS) >60.0, RMSMappingQuality <5.0, MQRankSumTest (MQRankSum) <-5.0, ReadPositionRankSum (ReadPosRankSum) <-5.0 to retain reliable SNPs. The processed SNPs were further filtered using VCFtools

version 0.1.13 (Danecek et al., 2011) to keep those with a minimum coverage of 5x and represented in at least 80% of the individuals.

In order to investigate the relative effect of the different SNP filters, other less conservative VCFtools filtering settings such as a reduced genotyping rate of 50% or reduced depth requirement of 2x were used, and the relative increase in number of SNPs recovered for each species was recorded. For each species, the resulting VCF files were then annotated with the names and coordinates of the original targets using retabvcf.pl (Matz, 2019). The targets represented in each species and the number of SNPs per target were then extracted from the annotated VCF files (Appendix S4).

To assess the applicability of probes designed from *L. bulimoides* and other related pteropod species, the relationship between sequence divergence and number of SNPs recovered was investigated. The genetic divergence between *L. bulimoides* and each of the four other species was calculated from the branch lengths of a maximum likelihood (ML) phylogeny of pteropods based on transcriptome data (Peijnenburg et al., 2020). The number of SNPs recovered per species using the most conservative filtering settings (80% genotyping rate and 5x depth) was plotted against sequence divergence from *L. bulimoides* in R (R Core Team, 2017).

Availability of data and materials

The genomic assembly (NCBI accession: SWLX00000000, BioSample ID: SAMN11131519), and raw sequencing data of the target capture are available in NCBI Genbank, under BioProject PRJNA527191. The transcriptome is available in NCBI Genbank under the NCBI accession SRR10527256 (BioSample ID: SAMN13352221, BioProject: PRJNA591100). The list of *L. bulimoides* contigs with homology to biomineralisation proteins and set of 82-mer probes developed for *L. bulimoides* are included as Supporting File 1 and 2 (https://doi.org/10.1186/s1286 4-019-6372-z). The additional information supporting the conclusions of this article are included as appendices within the Supplementary Information File.

Funding

This research was supported by the Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to K.T.C.A.P. Fieldwork was also supported by NSF grants OCE-1029478 and OCE-1338959 (to Erica Goetze). The Atlantic Meridional Transect program is supported by the UK NERC National Capability funding.

Acknowledgements

We thank Erica Goetze, Nina Bednaršek, Alice Burridge, and Lisette Mekkes as well as the captains and crews of the research cruises for support and assistance with collecting zooplankton samples.

Supplementary material

Appendix S1 Draft genome statistics

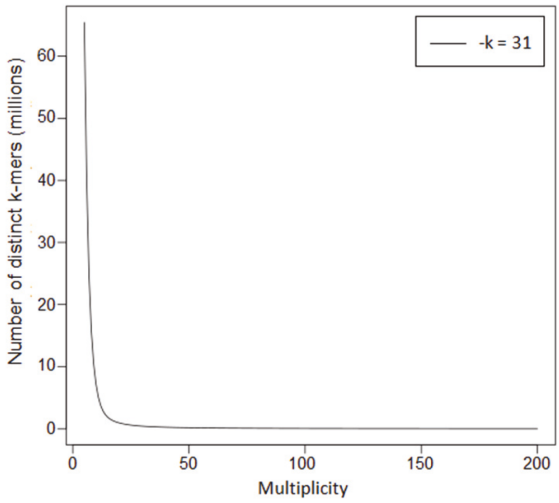


Figure S1 Histogram of k-mer frequency distribution in the assembled *Limacina bulimoides* draft genome for k = 31. The x-axis represents the number of times a k-mer occurred and the y-axis represents the number of distinct k-mers for the given multiplicity.

Appendix S2 Density of coverage per target for each species

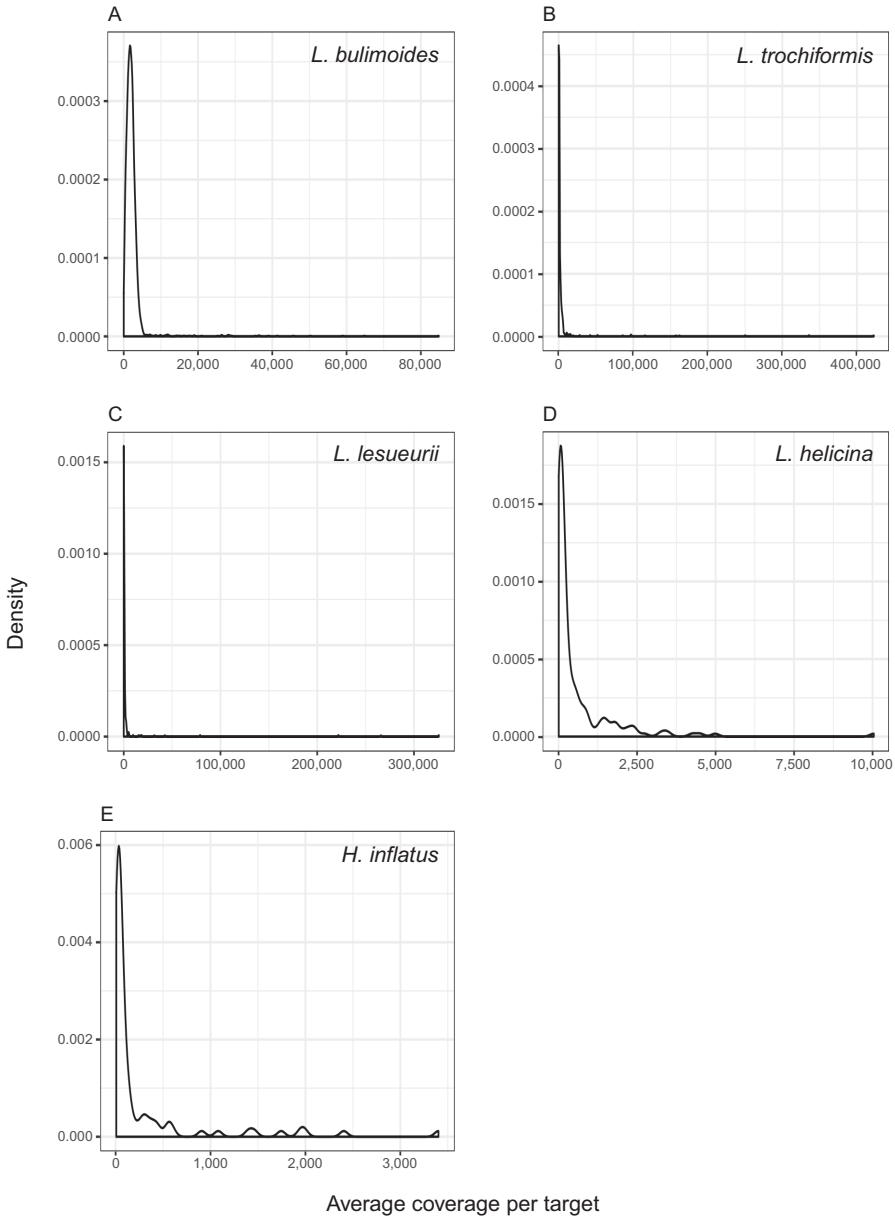


Figure S2 Density plot of coverage for each target, averaged across nine individuals, for each of the five shelled pteropod species (*Limacina bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. trochiformis*, *L. helicina* and *H. inflatus*).

Appendix S3 Specimen collection and molecular analyses

Draft genome

The sequenced *L. bulimoides* individual, Lbul_AMT22_57_08 was collected during the Atlantic Meridional Transect 22 (AMT22) cruise in November 2012 in the southern gyre of the Atlantic Ocean (25°44'S, 25°0'W). The specimen was photographed in a standardised orientation using a Zeiss V20 stacking microscope (Figure S1).

Modified DNA extraction protocol with E.Z.N.A. Insect Kit (Omega Bio-tek) for extracting DNA from shelled pteropods

1. Soak the complete individual in nuclease-free water for 30 minutes.

2. Transfer individual into 2 ml screw cap grinding tube containing:
- bashing beads
- 400 µl CTL buffer
- 28.5 µl proteinase K solution

3. Grind individual for 20 seconds at a frequency of 30 per second.

4. Incubate sample at 60 °C for 20 minutes.

5. Repeat step *3* and incubate sample at 60°C for 20 minutes.

6. Add 400 µl chloroform:isoamyl alcohol, vortex for 10 seconds.

7. Centrifuge at 13,500 g for 18 minutes at room temperature.

8. Prepare the HiBind DNA Mini Column:

Figure S3 Stacking microscopy photograph of the sequenced *Limacina bulimoides* specimen.

- assemble to collection tube
- add 100 µl 3 M NaOH (equilibration buffer) on filter
- incubate for 4 minutes and centrifuge at 17,000 g for 20 seconds.

9. Transfer upper aqueous phase (~320 µl) from step 7 to a new 1.5 ml eppendorf tube. Avoid the white milky surface.

10. Add one volume CBL buffer.

11. Add one volume 100% EtOH and vortex for 10 seconds.

12. Transfer 750 µl of the mixture to the equilibrated HiBind DNA Mini Column and centrifuge at 10,000 g for 1 minute. Discard the flow-through.

13. Repeat step 12 until all of the mixture has been used.

14. Warm-up the Elution Buffer at 70 °C.

15. Add 500 µl HBC buffer and centrifuge at 10,000 g for 30 seconds. Discard the flow-through.

16. Add 500 µl DNA Wash Buffer and centrifuge at 10,000 g for 1 minute. Discard the flow-through.

17. Repeat step 15.

18. Re-insert HiBind DNA Mini Column into new collection tube, centrifuge at 15,000 g for 2 minutes.

19. Re-insert HiBind DNA minicolumn into new eppendorf tube.

20. Add 40 µl of pre-heated Elution Buffer directly to the center of the Mini Column membrane without touching it.

21. Wait for 2 minutes and then centrifuge at 10,000 g for 1 minute.

22. Repeat steps 20 and 21 for a final elution volume of 80 µl.

Target capture protocol

All DNA libraries were fragmented to an average size of 300 bp by sonication and were prepared using the NEXTflex™ Rapid Pre-Capture Combo Kit (Bioo Scientific, Austin, TX, USA), including a step of single adapter indexing of each library. Libraries were cleaned-up, amplified separately for 8 and pooled in the following way: the nine *L. bulimoides* were part of a pooled capture that included 27 specimens of *L. bulimoides* in total. The four other species were amplified separately for another 8 cycles to increase the amount of DNA, and then combined into two pools that contained all four species with a total of 22-23 specimens per pool.

We increased the efficiency of the hybridisation and aimed to maximise the number of on-target captured sequences by doing the capture reaction twice, splitting the total amount of baits required for one reaction in two. The first round of hybridisation was performed using 4 µl of baits for each reaction. The reaction was performed over three days at a temperature of 60 °C in order to maximise the specificity. Capture was performed consecutively using DYNAbeads MyOne Streptavidin C1 beads (Invitrogen) to bind the hybridised targets during 30 min at 65 °C. The captured DNA was amplified by PCR for 8 cycles using KAPA HiFi HotStart

ReadyMix (Kapa Biosystems). A second round of hybridisation was conducted using 1.5 µL of baits for each of the two pools, followed by a second capture and 6 more cycles of post-capture PCR. Finally, the two mixed-species pools were mixed together in equal proportions and sequenced on a NextSeq 550 (Illumina) with a 2x150 bp mid-output kit v.2. The *L. bulimoides* pool was mixed equally with two other *L. bulimoides* pools containing 27 individuals and sequenced on a NextSeq 550 (Illumina) with a 2x150 bp high-output kit v.2.

Appendix S4 Commands used for SNP calling from fastq files

For manipulating targets fasta file: Used concatFasta.pl https://github.com/z0on/2bRAD_denovo/blob/master/concatFasta.pl (concatenate fasta file into user-specified number of "chromosomes") and retabvcf.pl from https://github.com/z0on/2bRAD_denovo/blob/master/retabvcf.pl (re-annotate .vcf file) to increase speed and reduce memory requirements. This is done before mapping reads to the targets.

(1) Use mapping.sh on demulitplexed, raw fastq.gz files

Contents of mapping.sh:

```
## mapping (input: *_R1.fastq.gz, *_R2.fastq.gz, <draft_genome_concatenated>.fasta)
## collect file names and place in file called names
ls *_R1.fastq.gz >names
sed -e s/_R1.fastq.gz//g -i names

## use parallel to run 24 threads of bwa mem, estimate two threads per specimen, for computer with 48 threads, producing aligned bam file *_aln.bam per specimen
cat names | parallel —verbose -j 24 "
bwa mem -M Genomic_targets_Limacina_bulimoides_cc.fasta {}_R1.fastq.gz {}_R2.fastq.gz | samtools view -Sbh - -o {}_aln.bam"

##cleaning initial bam files
cat names | parallel —verbose -j 10 "
samtools view {}_aln.bam | fgrep XA | cut -f 1 > bad_names_{}.txt
samtools view -h {}_aln.bam  | fgrep -vf bad_names_{}.txt | samtools view -Sb - > {}_aln2.bam
samtools view {}_aln2.bam | fgrep SA | cut -f 1 > bad_names_{}.txt
samtools view -h {}_aln2.bam | fgrep -vf bad_names_{}.txt | samtools view -Sb - > {}_aln3.bam
samtools view -b -F 3332 -f 3 {}_aln3.bam > {}_aln3_cleaned.bam"

##sort reads (3min)
chmod 755 *_aln3_cleaned.bam
cat names | parallel —verbose -j 10 "
samtools sort {}_aln3_cleaned.bam -o {}_sorted.bam"

##mark duplicates
cat names | parallel —verbose -j 10 "
java -jar picard.jar MarkDuplicates I={}_sorted.bam O={}_dedup.bam M={}_dedup_metricsfile ASSUME_SORT_ORDER=coordinate VALIDATION_STRINGENCY=SILENT REMOVE_DUPLICATES=true &>Log_{}_dedup.txt"
```

```
##Add or replace read groups, most important is RGSM which allows GATK to call genotypes by individuals
cat names | parallel —verbose -j 15 “
java -jar picard.jar AddOrReplaceReadGroups I={}_dedup.bam O={}_dedup_RG.bam SORT_ORDER=coordinate RGLB=lib1 RGPL=illumina RGPU=unit54 RGSM={}”

##index .bam file
for file in *_RG.bam; do
samtools index $file “${file}.bai”;
done

#end of mapping.sh
```

(2) The steps below follow Best Practices for GATK3.8., for genotyping SNPs from bam files

```
##HaplotypeCaller, use gnu-parallel to use multiple cores
cat names | parallel —verbose -j 20 “java -jar GenomeAnalysisTK.jar \
-R Genomic_targets_Limacina_bulimoides_cc.fasta \
-T HaplotypeCaller \
-I {}_dedup_RG.bam \
—emitRefConfidence GVCF \
-o $src/gvcf/{}.g.vcf”

##CombineGVCFs- files are combined per population at this step, so parallel is no longer needed.
java -jar GenomeAnalysisTK.jar \
-R ../Genomic_targets_Limacina_bulimoides_cc.fasta \
-T CombineGVCFs \
—variant Ltro.list \
-o Ltro.g.vcf

##GenotypeGVCFs
java -jar GenomeAnalysisTK.jar
-R ../Genomic_targets_Limacina_bulimoides_cc.fasta \
-T GenotypeGVCFs \
—variant Ltro.g.vcf \
—max_alternate_alleles 18 \ #depends on organism (diploid) and number of individuals (n=9)
-o ../combined/Ltro_raw_snps.vcf

##SelectVariants
for file in *.vcf; do
```

```
java -jar GenomeAnalysisTK.jar \
- T SelectVariants \
- R ../Genomic_targets_Limacina_bulimoides_cc.fasta \
- V $file\
- selectType SNP \
- o ${file%_variants*}_snps.vcf;
done

##VariantFiltration
for file in *_raw_snps.vcf; do
java -jar GenomeAnalysisTK.jar \
-T VariantFiltration \
-R  ../Genomic_targets_Limacina_bulimoides_cc.fasta \
-V $file \
—filterExpression “QD < 2.0 || FS > 60.0 || MQ < 50.0 || MQRankSum < -5.0 ||
ReadPosRankSum < -5.0” \
—filterName “choquet2018” \
-o ${file%raw*}filtered_snps.vcf &> ${file%raw*}filtered_snps.log ;
done
```

(3) The steps below use vcftools to filter the snps within the vcf file and retabvcf.pl to rename coordinates of concatenated contigs with the names of the targets

```
##Filter SNPs
for file in *filtered_snps.vcf; do
vcftools —vcf $file —max-missing 0.8 —min-meanDP 5.0 —recode —recode-INFO-
all —out ${file%filtered*}cleaned_80_minDP5;
done

##Rename “chromosomes” in vcf file with name of the targets
for file in *_cleaned_80_minDP5.recode.vcf; do
~/2bRAD_denovo/retabvcf.pl vcf=$file tab=../Genomic_targets_Limacina_
bulimoides_cc.tab >retab_$file;
done

##To count number of targets present in the vcf file
for file in retab_*; do
awk ‘/#CHROM/,0’ $file |cut -f1 |uniq -c > ${file%cleaned*}target_uniqc.list;
done
```

Appendix S5 Number of SNPs per target (individual plots)

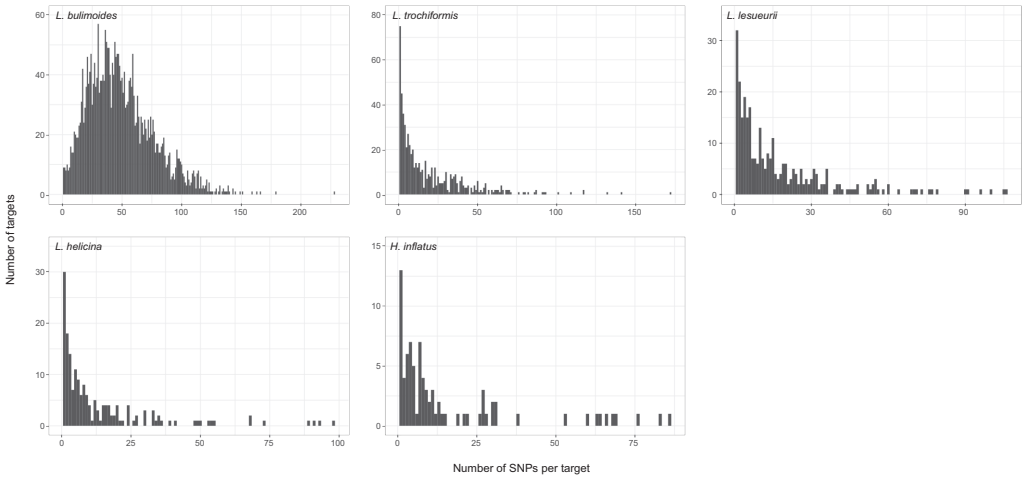


Figure S4 Number of SNPs per recovered target for each of the five shelled pteropod species (*Limacina bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. helicina* and *Heliconoides inflatus*), based on filtering settings of minimum presence in 80% of individuals with at least 5x read depth.

Table S1 Number of targets with a minimum coverage of 15x in at least one base, across the five shelled pteropod species *Limacina bulimoides*, *L. trochiformis*, *L. lesueurii*, *L. helicina* and *Heliconoides inflatus*. The number of targets with more than 90% or 50%, and less than 10% of bases with 15x coverage is also displayed for each species.

Species	Total targets recovered at 15x depth or more	≥90% bases recovered with 15x depth or more	≥50% bases recovered with 15x depth or more	≤10% bases recovered with 15x depth or more
L. bulimoides	2822	2446	2768	5
L. trochiformis	620	42	206	88
L. lesueurii	302	16	72	77
L. helicina	177	6	15	64
H. inflatus	83	5	13	44

Table S2 Sampling localities and raw sequencing results per specimen.

Individual	RMNH number	NCBI	Latitude	Longitude	Raw reads	Final mapped reads	% HQ reads	% targets covered	Depth
Lbul_AMT24_20_03	RMNH.MOL.341197	SAMN11131474	-18.32	-25.08	9277276	4294018	46.29	97.52	311.44
Lbul_AMT24_20_04	RMNH.MOL.341198	SAMN11131475	-18.32	-25.08	4997100	1044990	20.91	96.50	74.07
Lbul_AMT24_20_09	RMNH.MOL.341199	SAMN11131476	-18.32	-25.08	9528311	3566992	37.44	97.40	251.49
Lbul_AMT24_22_04	RMNH.MOL.341200	SAMN11131477	-24.45	-25.05	17168774	3718492	21.66	97.52	263.53
Lbul_AMT24_22_06	RMNH.MOL.341201	SAMN11131478	-24.45	-25.05	6298310	1677720	26.64	96.94	113.11
Lbul_AMT24_22_08	RMNH.MOL.341202	SAMN11131479	-24.45	-25.05	8394293	3357164	39.99	97.19	238.12
Lbul_AMT24_23_06	RMNH.MOL.341203	SAMN11131480	-27.77	-25.02	10546514	2941122	27.89	97.77	204.06
Lbul_AMT24_23_08	RMNH.MOL.341204	SAMN11131481	-27.77	-25.02	14110230	5391732	38.21	97.80	380.22
Lbul_AMT24_23_09	RMNH.MOL.3401205	SAMN11131482	-27.77	-25.02	14440677	5786036	40.07	97.57	410.24
Ltro_AMT24_19_01	RMNH.MOL.340274	SAMN11131501	-14.67	-25.07	19439184	2605782	13.40	22.75	693.79
Ltro_AMT24_19_02	RMNH.MOL.340275	SAMN11131502	-14.67	-25.07	12770462	1741952	13.64	21.42	453.24
Ltro_AMT24_19_03	RMNH.MOL.340276	SAMN11131503	-14.67	-25.07	13776593	1783082	12.94	20.23	474.98
Ltro_AMT24_20_02	RMNH.MOL.340277	SAMN11131504	-18.32	-25.08	9152848	935860	10.22	19.40	238.24
Ltro_AMT24_20_03	RMNH.MOL.340278	SAMN11131505	-18.32	-25.08	12020467	1979350	16.47	20.47	525.52
Ltro_AMT24_20_05	RMNH.MOL.340279	SAMN11131506	-18.32	-25.08	11296879	1034722	9.16	17.13	283.11
Ltro_KOK1703_03_01	RMNH.MOL.340280	SAMN11131507	22.65	-157.69	16580239	1664588	10.04	22.03	408.59
Ltro_KOK1703_03_02	RMNH.MOL.340281	SAMN11131508	22.65	-157.69	22810256	2082476	9.13	19.53	571.05
Ltro_KOK1703_03_03	RMNH.MOL.341274	SAMN11131509	22.65	-157.69	21721894	2059844	9.48	19.91	564.29
Lles_AMT22_25_01	RMNH.MOL.341206	SAMN11131483	20.40	-38.61	5422102	894512	16.50	12.78	505.10
Lles_AMT22_25_03	RMNH.MOL.341207	SAMN11131484	20.40	-38.61	7054061	694824	9.85	14.40	379.47
Lles_AMT22_25_04	RMNH.MOL.340258	SAMN11131485	20.40	-38.61	2998157	468110	15.61	11.98	269.37
Lles_AMT22_49_01	RMNH.MOL.340259	SAMN11131486	-15.30	-25.07	6755374	886610	13.12	15.59	485.75
Lles_AMT22_49_03	RMNH.MOL.340260	SAMN11131487	-15.30	-25.07	6942361	718982	10.36	11.03	414.07
Lles_AMT22_49_04	RMNH.MOL.340261	SAMN11131488	-15.30	-25.07	6789274	684160	10.08	10.12	405.38
Lles_AMT22_55_02	RMNH.MOL.340262	SAMN11131489	-24.13	-25.00	9641976	988170	10.25	14.77	522.79
Lles_AMT22_55_03	RMNH.MOL.340263	SAMN11131490	-24.13	-25.00	8761780	1139976	13.01	15.53	449.81
Lles_AMT22_55_04	RMNH.MOL.340264	SAMN11131491	-24.13	-25.00	9170668	787034	8.58	13.31	448.47
Lhel_AMT24_27_02	RMNH.MOL.340265	SAMN11131492	-40.12	-30.92	12497407	486658	3.89	14.18	95.55
Lhel_AMT24_27_03	RMNH.MOL.340266	SAMN11131493	-40.12	-30.92	9928760	398788	4.02	13.50	71.00
Lhel_AMT24_27_04	RMNH.MOL.340267	SAMN11131494	-40.12	-30.92	12041919	365014	3.03	14.66	65.82
Lhel_AMT24_28_03	RMNH.MOL.340268	SAMN11131495	-41.48	-33.87	21381584	621646	2.91	16.72	105.40
Lhel_AMT24_28_04	RMNH.MOL.340269	SAMN11131496	-41.48	-33.87	12948171	391250	3.02	13.51	69.07
Lhel_AMT24_28_05	RMNH.MOL.340270	SAMN11131497	-41.48	-33.87	14792208	407548	2.76	12.24	64.54
Lhel_CCE_123_02	RMNH.MOL.340271	SAMN11131498	48.36	-126.31	4010923	165092	4.12	10.81	43.32
Lhel_CCE_123_03	RMNH.MOL.340272	SAMN11131499	48.36	-126.31	3787224	129550	3.42	9.26	35.30

TABLE S2 Continued

Individual	RMNH number	NCBI	Latitude	Longitude	Raw reads	Final mapped reads	% HQ reads	% targets covered	Depth
Lhel_CCE_123_06	RMNH.MOL.340273	SAMN11131500	48.36	-126.31	1721637	70278	4.08	8.23	23.13
Hinf_AMT22_21_02	RMNH.MOL.340283	SAMN11131510	25.48	-39.00	3890761	92610	2.38	9.99	40.60
Hinf_AMT22_21_03	RMNH.MOL.340284	SAMN11131511	25.48	-39.00	3469925	55336	1.59	6.16	22.73
Hinf_AMT22_21_06	RMNH.MOL.340285	SAMN11131512	25.48	-39.00	1609269	31112	1.93	5.13	23.76
Hinf_AMT22_45_01	RMNH.MOL.340286	SAMN11131513	-8.08	-25.04	3726617	77294	2.07	12.80	34.87
Hinf_AMT22_45_02	RMNH.MOL.340287	SAMN11131514	-8.08	-25.04	3920742	91946	2.35	12.58	55.57
Hinf_AMT22_45_03	RMNH.MOL.340288	SAMN11131515	-8.08	-25.04	1256559	24332	1.94	4.16	13.01
Hinf_AMT22_66_01	RMNH.MOL.340289	SAMN11131516	-38.08	-39.31	4229048	107418	2.54	10.85	52.23
Hinf_AMT22_66_02	RMNH.MOL.340290	SAMN11131517	-38.08	-39.31	3653019	71554	1.96	6.28	22.93
Hinf_AMT22_66_04	RMNH.MOL.340291	SAMN11131518	-38.08	-39.31	2042221	38566	1.89	5.89	20.99

4

Genome-wide phylogeography reveals cryptic speciation in a circumglobal planktonic calcifier: *Limacina bulimoides* (Pteropoda)

L.Q. Choo, G. Spagliardi, M. Choquet, E. Goetze, G. Hoarau, K.T.C.A. Peijnenburg

Abstract

One of the big questions in marine speciation is how species originate without apparent barriers to gene flow. By studying the phylogeography of planktonic species, we can uncover processes and evolutionary forces that facilitate speciation in the open ocean. Shelled pteropods possess thin aragonitic shells susceptible to dissolution and much research has focused on estimating the sensitivity of *Limacina* species to ocean acidification. It is thus of high priority to define their species boundaries, as distinct evolutionary units are likely to respond differently to changing ocean conditions. In this first global population genomics study of a marine zooplankton species, we assessed spatial patterns of divergence across 161 individuals of the pteropod *Limacina bulimoides* with 107,214 single nucleotide polymorphisms and morphological variation, to identify evolutionary drivers of divergence. Clustering and admixture analyses revealed three distinct genetic lineages with at least 30% genetic differentiation (pairwise F_{ST}) between them: one Atlantic lineage, sampled across the Atlantic Ocean, an Indo-Pacific lineage, sampled in the North Pacific and Indian Ocean, and a Pacific lineage, sampled in the North and South Pacific. No recent gene flow occurred among lineages, even in areas of sympatry in the North Pacific, and further geographic structuring occurs within all three lineages. A time-calibrated phylogeny suggests that the Atlantic and the Indo-Pacific/ Pacific lineages diverged about one million years (My) ago, with the highest effective population size (N_e) reconstructed during predominantly glacial periods, and a decrease in N_e following the start of the Holocene. The three main lineages are reproductively isolated, yet morphologically cryptic, with subtly different shell shape distributions and variable shell colour. While the circumglobal *L. bulimoides* species complex consists of smaller species ranges and N_e than initially thought, these pteropods still possess high levels of genetic variability and have thrived through repeated Pleistocene glacial transitions. Whether they can cope with the anticipated rate of future ocean change, however, remains to be seen.

Introduction

We still know little about how species arise and persist in the open ocean, a seemingly interconnected habitat with few apparent barriers to dispersal. Allopatric speciation was accepted as the dominant mode of speciation for a long time (Coyne and Orr, 2004; Mayr, 1954; Norris, 2000), but it has been increasingly documented that speciation can occur in the absence of barriers to gene flow, particularly in marine systems (Bierne et al., 2003; Endler, 1977; Johannesson et al., 2010; Potkamp and Fransen, 2019; Schluter, 2009). To gain insight into the mechanisms of speciation in marine habitats, it is important to understand the factors that facilitate or constrain divergence, rather than categorising case studies into allopatric, parapatric or sympatric speciation (Butlin et al., 2008; Fitzpatrick et al., 2009). Marine species have a wide variety of life history traits, dispersal potentials and extent of gene flow between populations, and holoplanktonic species represent an extreme end of this scale with high fecundities, enormous population sizes, and huge potential for dispersal and gene flow.

Marine zooplankton are among the most abundant multicellular eukaryotes in the world. With their unique life history traits, marine zooplankton are an interesting system to assess the impact of large population sizes and high dispersal potential on speciation and divergence (Bucklin et al., 2021a; Faria et al., 2021). The large effective population sizes in marine zooplankton can have varying effects on evolution depending on the relative effects of selection and stochastic demographic effects. There can be higher levels of genetic diversity for selection to act upon and more effective selection due to lower effects of random processes like genetic drift (Barrio et al., 2016; Peijnenburg and Goetze, 2013). However, the relationship between adaptation and population size may not be straightforward (Galtier, 2016), and the reduction of genetic drift can decrease opportunities for the accumulation of genetic differentiation between (partially) isolated populations, e.g., in mutation-order speciation (Mani and Clarke, 1990; Schluter, 2009).

With the increased availability of various powerful genetic tools, there has been growing evidence for cryptic species complexes in the open ocean, contrary to historical expectations that many marine pelagic species have circumglobal, homogeneous populations (Norris, 2000; van der Spoel and Heyman, 1983). Species with widespread geographical distributions can have slight morphological variation across their range that may be insufficient to delimit sibling species (Knowlton, 1993), or in some cases well-characterised differences are deemed as the result of a highly variable morphospecies (e.g., Apolônio Silva De Oliveira et al., 2017), resulting in an underestimation of ecologically relevant divergence with conventional morphology-based taxonomy (Bongaerts et al., 2021). A majority of circumglobal planktonic species have been found to harbour cryptic diversity based on genetic evidence (e.g., Andrews, Norton, Fernandez-Silva, Portner, & Goetze, 2014; Burridge, Van Der Hulst, Goetze, & Peijnenburg, 2019; Casteleyn et al., 2010; Cornils, Wend-Heckmann, & Held, 2017; Halbert, Goetze, & Carlon,

2013; Hirai, Tsuda, & Goetze, 2015; Miyamoto, Machida, & Nishida, 2012; Wall-Palmer et al., 2018; Whittaker and Rynearson, 2017). An integrative taxonomy framework (Burridge et al., 2019; Padial, Miralles, De la Riva, & Vences, 2010) would be beneficial to identify consistent morphological differences between these lineages and support the current genetic findings (McManus and Katz, 2009). Morphological variations congruent with genetic clines can indicate ecological selection or phenotypic plasticity, while the absence of morphological divergence between genetically distinct populations can result from recent colonisation, widespread gene flow or stabilising selection (Fišer et al., 2018; Milá et al., 2017).

Pteropods play important ecological and biogeochemical roles globally (Buitenhuis et al., 2019; Fabry et al., 2009; Hunt et al., 2008; Manno et al., 2010; Sulpis et al., 2021) and are commonly used as bioindicators for the effects of ocean acidification on marine calcifiers (Bednaršek, Harvey, Kaplan, Feely, & Možina, 2016; Fabry, McClintock, Mathis, & Grebmeier, 2009; Manno et al., 2017). However, little is known about their adaptive potential for future environmental changes, including their levels of genetic diversity and functional genomic variation. Pteropods possess thin aragonitic shells, which are vulnerable to dissolution and harder to make in acidified waters as predicted for the future (Busch et al., 2014; Mekkes et al., 2021b) and as observed along contemporary gradients of ocean acidification (Bednaršek et al., 2018; Manno et al., 2018; Mekkes et al., 2021a; Niemi et al., 2021). Shelled pteropods feed on phytoplankton and particulate matter, including bacteria and small protists, by trapping them in external mucous webs and ingesting the webs (Conley, Lombard, & Sutherland, 2018; Hunt et al., 2008; Lalli and Gilmer, 1989). Pteropods constitute a significant part of the diet of fish and other predators from higher trophic levels in the pelagic food web (Groot and Margolis, 1991; Hunt et al., 2008; Lalli and Gilmer, 1989). An important step for understanding the capacity of species to adapt to future environmental changes is to assess the spatial distribution of genetic variation and potential for gene flow across their species range (Bell, 2013; Harvey et al., 2014; Manno et al., 2017; Munday et al., 2013; Poloczanska et al., 2016; Sunday et al., 2014).The aragonitic shell of pteropods has traditionally been used to classify and identify species of pteropods (Bé and Gilmer, 1977; Lalli and Gilmer, 1989; van der Spoel et al., 1997). The use of shell shape as a diagnostic character for species identification has allowed resolving ancient species boundaries due to their traceable morphology preserved in fossils (Janssen and Peijnenburg, 2017). However, a solely morphological approach to extant pteropod taxonomy was shown to be inadequate, with the identification of genetically distinct cryptic species within described morphospecies (Hunt et al., 2010; Jennings, Bucklin, Ossenbrügger, & Hopcroft, 2010; Maas, Blanco-Bercial, & Lawson, 2013). More recent case studies have used both genetic data from barcoding genes and geometric morphometrics of the shell shape to assess species boundaries with good success (e.g., Burridge et al., 2015a, 2019; Choo et al., 2021; Shimizu et al., 2021).

Among shelled pteropods, species of the *Limacina* genus are protandrous hermaphrodites, developing from larvae into males first, before transitioning into females that lay free floating egg strings from which free-swimming veliger larvae hatch (Lalli and Wells, 1978). *Limacina bulimoides* is the most common warm water *Limacina* species worldwide, and is found in all tropical and subtropical oceans from ~45°N to ~40°S (Bé and Gilmer, 1977). This species is estimated to complete their life cycle in less than one year (Wells, 1976). Our prior work on *L. bulimoides* based on two barcoding genes and shell shape differences (Choo et al., 2021) revealed two dispersal barriers in the Atlantic Ocean: one in the southern subtropical gyre and the other across the equatorial upwelling region (Choo et al., 2021). However, genome-wide loci and observations over a global spatial scale are required to assess differences in genomic variation across globally-distributed zooplankton populations (Bucklin et al., 2018; Choquet et al., 2019) and to detect signals of incipient speciation on multiple isolated and physically unlinked genes (Feder and Nosil, 2010).

To identify the origins and drivers of dispersal barriers in the open ocean, we conducted the first global population genomics study of a marine zooplankton species. We investigated the evolutionary history and extent of reproductive isolation between independently diverging lineages within the pteropod species, *L. bulimoides*, with genome-wide single nucleotide polymorphisms (SNPs) from target capture loci characterized in (Choo et al 2020). We also measured phenotypic differences including shell shape and examined these within the geographic and genetic context of *L. bulimoides*, and explored historical drivers of speciation in the tropical and subtropical ocean.

Material and Methods

Sample collection

Bulk plankton samples were collected in the Atlantic, Pacific and Indian Oceans during several cruises (Table 1). Specimens were collected by using either oblique or horizontal tows with ring nets or bongo nets (mesh size 200-505µm). Tows were conducted for 20 minutes to 1 hour, at approximate depths ranging between 60 and 370 m. Samples were immediately preserved in 96% ethanol and stored at -20 °C. The ethanol was replaced after 24 hours of preservation. *Limacina bulimoides* specimens were subsequently sorted in the laboratory and photographed in a standard orientation prior to destructive genetic work.

Library prep and sequencing

Specimens were prepared for sequencing as described in (Choo et al., 2020). In brief, genomic DNA was extracted from each individual with either the E.Z.N.A mollusc or insect DNA extraction kit (Omega Bio-Tek) (Supplementary Material 2.1). The DNA was sheared by sonication to attain a peak length of 300 bp with a Covaris S220 or ME220 focused ultrasonicator. After sonication, the fragmented DNA was

TABLE 1 Collection details of specimens of *Limacina bulimoides* from the global ocean. Number of specimens per sampling location used in target capture and geometric morphometric analyses are indicated. The collection dates are in the format DD/MM/YYYY. See also map in FIGURE 1a.

Cruise and station	Latitude	Longitude	Date collected	Ocean basin	Target capture	Morphometrics
NIC2_S1C3	21°16'N	21°01'W	31/12/2017	N. Atlantic	11	11
NIC8_S5C3	38°45'N	56°18' W	14/4/2018	N. Atlantic	11	11
AMT24_17	7°28'S	25°07'W	15/10/2014	Eq. Atlantic	11	11
NIC2_S9C3	5°58'N	47°50'W	13/1/2018	Eq. Atlantic	10	10
AMT24_22	24°27'S	25°03'W	21/10/2014	S. Atlantic	10	10
AMT24_23	27°46'S	25°01'W	22/10/2014	S. Atlantic	10	10
SN105_08	4°23'E	67°00'E	10/12/2015	Indian	11	10
KOK1703_03	22°39'N	157°41'W	3/7/2017	N. Pacific	10	10
KH1110_02	23°00'N	160°00'E	7/12/2011	N. Pacific	11	11
KH1110_05	23°00'N	179°59'E	14/12/2011	N. Pacific	11	11
KH1110_08	22°47'N	158°06'W	19/12/2011	N. Pacific	11	10
KH1110_15	23°00'S	119°15'W	8/1/2012	S. Pacific	11	11
KH1110_18	29°59'S	107°00'W	13/1/2012	S. Pacific	11	11
KH1110_21	23°00'S	100°00'W	18/1/2012	S. Pacific	11	11
SO255_143	32°52'S	179°46'W	3/4/2017	S. Pacific	11	11
Total					161	159

prepared into individual libraries using the NEXTflex Rapid Pre-Capture Combo Kit (Bioo Scientific). Individually barcoded libraries were subsequently pooled at equimolar concentrations with 26 to 27 libraries per pool. The target capture reaction was then performed on each pool using the myBaits Custom Target Capture kit (Arbor Biosciences), with a capture probe set specifically designed for *L. bulimoides*, as described in Choo et al. (2020). The baits target 2890 nuclear regions as well as the mitochondrial cytochrome *c* oxidase subunit I (COI) and nine other mitochondrial gene regions. To maximize the specificity of the capture reaction, the hybridisation time of the probes with the libraries was extended to three days, and the capture was performed twice, with 4 µl and 1.5 µl of probe mix, respectively (Choo et al., 2020) (SUPPLEMENTARY MATERIAL 2.2). Captured library pools were sequenced in paired-end on the Illumina NextSeq 500 platform using three high-output v2 chips (150 cycles).

QUALITY FILTERING

Raw sequences were de-multiplexed (Genbank accessions: SAMN11131477-79, SAMN11131480-82, SAMN20293115-269), and then mapped with BWA mem 0.7.12 (Li, 2013) to the reduced contig set of the genomic assembly (Genbank accession: SWLX00000000). This reduced genomic assembly contains all the contigs that were included for the probe design. The resulting alignments were cleaned and filtered with SAMtools version 1.4.1 (Li et al., 2009) to retain only properly and uniquely mapped pairs. Duplicates were marked and removed with Picard version 2.18.5 (http://broadinstitute.github.io/picard). Variant calling was done using GATK

4.1.7.0, following the Variant Discovery Pipeline (Auwera et al., 2013; Depristo et al., 2011) with GNU parallel utility (Tange, 2011). Haplotypes were called individually using HaplotypeCaller with emitRefConfidence output with and without the setting EMIT_ALL_CONFIDENT_SITES in two separate instances. The resulting gVCF files for each setting were combined with CombineGVCFs. The combined gVCF files were then genotyped using GenotypeGVCFs. Single nucleotide polymorphisms (SNPs) were then extracted from the total variants using SelectVariants (-SelectType SNP). SNPs were hard-filtered with VariantFiltration using QualByDepth (QD) <2.0, FisherStrand (FS) >60.0, RMSMappingQuality <5.0, MQRankSumTest (MQRankSum) <-5.0, ReadPositionRankSum (ReadPosRankSum) <-5.0 to retain reliable SNPs. The SNPs were subsequently processed in BCFtools version 1.7 (https://github.com/samtools/bcftools) to retain only high-quality SNPs (with Phred score >20) covered at least three times and present in at least 80% of the genotypes (Supplementary Material 3.1).

Population structure

Patterns of population structure were assessed on the filtered dataset using only nuclear SNPs, after removing SNPs mapped to known mitochondrial contigs from the dataset. This was done to ensure that there was no conflict between the nuclear and mitochondrial SNPs, which could have distinct evolutionary histories. To visualise population structure, we conducted a Principal Component Analysis (PCA) in PLINK v1.90b6.17 (Chang et al., 2015), on a subset of independent SNPs, using linkage pruning settings of window size of 50 genetic variants, window shift of 10 variant counts and r^2 of 0.2. In addition, we used Admixture (Alexander et al., 2009) to infer populations and ancestries, with the same filtered and unlinked SNP dataset as used for PLINK. The analysis was conducted for values of K ranging between 2 and 7, and the cross validation error for each K value was calculated to determine the number of putative lineages.

In addition to clustering analyses, we investigated population structure based on nearest neighbour haplotype co-ancestry with fineRADstructure v.0.3.2 (Malinsky et al., 2018). We used the hapsFromVCF function of RADpainter to convert a biallelic filtered dataset of 96,355 nuclear SNPs into input format with SNPs from each of the 2609 contigs condensed into haplotypes. A locus was considered missing if not sequenced in more than half of the individuals. Of the 161 individuals, five were excluded as they had a higher proportion (>4%) of loci which were marked as missing. After excluding these individuals, we used the paint function of RADpainter to calculate a co-ancestry matrix, which summarises the nearest neighbour haplotype relationships in the dataset. Next, a clustering dendrogram of shared ancestry was inferred from the co-ancestry matrix using the fineSTRUCTURE Markov chain Monte Carlo clustering algorithm, with 100,000 burn-in iterations, 100,000 sample iterations, and thinning of 1000. The inferred clusters were arranged with a simple tree building algorithm in fineSTRUCTURE with 10,000 hill-climbing iterations.

To estimate the extent of population substructure, global weighted F_{ST} estimates for each lineage was calculated from the SNPs previously filtered for quality and with an additional filter of minor allele frequency (MAF) <0.01. To avoid bias due to genetic linkage of some loci, distribution of global weighted F_{ST} was estimated for each lineage with thinned SNP datasets, where one SNP is randomly chosen per contig and weighted F_{ST} is calculated for each thinning iteration for 1000 iterations (achieved using a custom script from Choquet et al., 2019). The resulting distributions for each lineage were plotted in R (R Core Team, 2017).

Genetic signals from the targeted mitochondrial COI fragments were separately analysed in a haplotype network after mapping and *de novo* assembly in Geneious Prime 2021.1.1 (https://www.geneious.com). First, all raw Illumina NextSeq reads of each of the 161 individuals were mapped to a reference database containing 355 unique *L. bulimoides* COI sequences of 564 base pairs sequenced from the Atlantic and Pacific Ocean (NCBI: MN952611-MN952965), using the Geneious assembler with medium-low sensitivity. The mapped reads were extracted and *de novo* assembled per individual using the Geneious assembler with medium sensitivity. The resulting contigs were annotated for the COI region, and the COI annotation (564 bp) with the highest coverage for each individual was translated to check that there were no stop codons or frameshift mutations before being extracted. The 161 COI annotations (Genbank accessions: MZ542566 - MZ542726) were combined with one representative from each known haplogroup from (Choo et al., 2021): the Atlantic haplogroups 1 and 2, and the Pacific haplogroup (Genbank accessions: MN952611, MN952937, MN952944), and used to create a multiple sequence alignment using MAFFT v7.222 (Katoh and Standley, 2013). A minimum spanning network was calculated from the alignment and visualised in POPART (Leigh and Bryant, 2015).

Genetic diversity

To estimate genetic diversity within lineages, we calculated heterozygosity and nucleotide diversity. Heterozygosity was calculated from the previously filtered SNP dataset in VCFtools 0.1.15 (Danecek et al., 2011) with the setting –het, which outputs a measure of heterozygosity on a per-individual basis. We plotted these values according to their lineage to obtain the lineage-specific distribution of heterozygosity.

For nucleotide diversity, equivalent filtering settings were used on a VCF dataset including both variant and non-variant sites (HaplotypeCaller —output-mode EMIT_ALL_CONFIDENT_SITES). Nucleotide diversity was calculated in VCFtools 0.1.15 (Danecek et al., 2011) for each genetic lineage, with the SNPs partitioned into their respective lineages, in windows of 2865 bp corresponding to the average contig size. Distributions of heterozygosity and nucleotide diversity for each genetic lineage were plotted in R version 4.0.3 (R Core Team, 2017) and a one-way ANOVA was used to test for significant differences between lineages.

Phylogenetic and demographic analyses

To infer a time-calibrated Bayesian tree, we generated a new SNP dataset. SNP calling was performed with all 161 *L. bulimoides* individuals and 11 *L. trochiformis* individuals (NCBI: SAMN111131501-09, SAMN20293270, SAMN20293271 from Choo et al., 2020). The *L. trochiformis* individuals were included to root the phylogeny and calibrate the age of divergence. SNP calling was followed by quality filtering using settings described in section 2.3 to obtain a dataset of biallelic nuclear SNPs. To reduce computational effort, we randomly selected two *L. trochiformis* individuals and 10 individuals from each of the *L. bulimoides* genetic lineages to include in our analyses. SNPs corresponding to these 32 selected individuals were re-filtered to retain only biallelic nuclear SNPs. This resulting SNP dataset was thinned with VCFtools 0.1.15 (Danecek et al., 2011) to select 1 SNP every 100 bp to reduce linkage between the SNPs, resulting in a final dataset of 19,669 thinned, biallelic nuclear SNPs. A SNAPP (Bryant et al., 2012) species tree was inferred in BEAST2 v2.6.2 (Bouckaert et al., 2019), using the approach demonstrated in (Stange et al., 2018). We used a custom ruby script (https://github.com/mmatschiner/snapp_prep/blob/master/snapp_prep.rb), with a strict clock substitution rate calibrated with an *a priori* age constraint of divergence between *L. trochiformis* and *L. bulimoides* (Peijnenburg et al., 2020). The script was modified to allow for unlinked population sizes between the three lineages. The divergence time between *L. trochiformis* and *L. bulimoides* was estimated at 13.57 Ma (95% confidence interval (CI) = 9.2804-20.9182 Ma), which was approximated to a lognormal distribution of (0, 13.57, 0.2), with a 95% CI of 8.99-19.7. The generation time of *L. bulimoides* was assumed to be one year in order to calibrate the estimated effective population size. We used Tracer (Rambaut et al., 2018) to observe that all Effective Sample Size (ESS) parameters converged to stationarity with values above 200 after 1,500,000 MCMC iterations, for three independent trials. The maximum clade credibility trees with mean heights were produced in TreeAnnotator v2.6.0 (Bouckaert et al., 2014). The effective population size (N_e) and their 95% highest posterior density (HPD) confidence intervals for each of the three genetic lineages were calculated with θ from the SNAPP analysis, according to the following formula: $N_e = \theta/4\mu$, where μ, the mutation rate was given as the log_clock_rate divided by $1x10^6$ with a generation time of 1 year.

Demographic changes for each lineage were reconstructed using Stairway Plot v2.1.1 (Liu and Fu, 2020). We used biallelic SNPs covered at least ten times and with a missingness of less than 50%, and subsequently applied a per-individual sequencing depth filter, where all SNPs within 0.5 to 2 times the mean depth were retained. Using VCFtools 0.1.15 (Danecek et al., 2011), we removed sites that showed a heterozygosity excess across all individuals (P_HET_EXCESS<0.01). In order to perform a site frequency spectrum (SFS) analysis, which assumes unlinked sites, we pruned the dataset for linkage disequilibrium, with PLINK v1.90b6.17 (Chang et al., 2015), in windows of 50 genetic variants, with a window shift of 10

variants and a value of r^2 set to 0.2. The resulting VCF file was used to calculate the SFS for each lineage with easySFS (https://github.com/isaacovercast/easySFS) with downsampling of sequences to maximise the number of SNPs. After downsampling, the number of surveyed haploid sequences for each lineage was 56, 40 and 64 for the Atlantic, Indo-Pacific and Pacific lineages, respectively. The total number of confidently called sites (variable and fixed) included in the SFS for each lineage was 2,064,717. The mutation rate, μ, was estimated to be 0.00599 from the SNAPP analysis. For each SFS, 200 subsampling iterations were conducted, as in the default recommendations. Population size estimates were truncated at 100 years as there were insufficient samples to reconstruct demographic changes in more recent time. The timing of demographic changes was compared to the Marine Isotope Stages derived from Lisiecki and Raymo (2005).

Shell morphology

All 161 *L. bulimoides* individuals used for the genetic analyses above were photographed in a standardised apertural orientation using a Zeiss V20 stacking stereomicroscope with Axiovision software prior to destructive DNA extraction. The shell images were digitised at 11 (semi-) landmarks (Supplementary Material 3.2) using TpsUtil and TpsDig (Rohlf, 2015). 159 of the 161 *L. bulimoides* individuals had the complete set of landmarks and could be included in the geometric morphometric analysis. Coordinates of the (semi-) landmarks were analysed in TpsRelW (Rohlf, 2015), using a generalised least-squares Procrustes superimposition (Rohlf and Slice, 1990; Zelditch, Swiderski, Sheets, & Fink, 2004).

A repeatability analysis was conducted with a subset of 30 individuals, comprising two individuals from each of the 15 locations. Images of these individuals were landmarked at the 11 (semi-)landmarks by two independent researchers. Centroid size and relative warp (RW) scores between the pairs of images per specimen after Procrustes Fit were compared using intra-class coefficient (ICC) in PAST3.0 (Hammer et al., 2001), and ICC values greater than 0.75 were considered sufficiently repeatable.

We tested for significant variation in shell shape across genetic lineages with a non-parametric Permutational Multivariate Analysis of Variance (one-way PERMANOVA) using Euclidean distances and 9999 permutations with the vegan package in R (Oksanen et al., 2019). Only the six repeatable RWs were used in the one-way PERMANOVA. The first two RW axes were plotted to visualise shell shape variation for different genetic lineages of *L. bulimoides*. Additionally, a Canonical Variates Analysis (CVA) was conducted in R (R Core Team, 2017) to discriminate shell morphometric differences between genetic lineages. A one-way ANOVA with a post-hoc Tukey HSD test was also conducted in R to test if the means of the canonical variate for each genetic lineage were different from the other groups. We also examined the effect of sampling location on shell shape, by conducting the above analyses with the four ocean basins (Atlantic, Indian, N. Pacific and S. Pacific) instead of the three genetic lineages.

We also assessed shell colour variation by qualitatively scoring aperture colour as either transparent, pink, tan or red-brown. We chose to score the colour of the aperture as it was not obscured by the tissue, which could interfere with colour identification. Scoring was done by two independent researchers to ensure repeatability of the colour scores. We tested then whether aperture colour was randomly distributed across location using a Fisher's exact test of independence.

Results

Data processing

Across 161 *Limacina bulimoides* individuals, a mean of 10,122,331 raw reads were obtained per individual after sequencing. On average, 3,961,477 reads per individual were available for variant calling, after mapping. 39.8% of the reads were mapped with high quality, and 94.4% of these reads were on target, resulting in a global reads-on-target of 37.6% (Supplementary Material 1). The total number of SNPs obtained for all 161 individuals was 1,790,328, which was slightly reduced to 1,629,382 upon hard filtering. Most of the SNPs (1,246,720 and 328,298, respectively) could be located back to 2,089 of the 2169 targeted coding regions and the 633 of the 643 conserved protein coding orthologous regions, respectively (Supplementary Material 4.1.1). There were 262 SNPs that could be mapped to two mitochondrial gene regions, the COI and 12S gene. Further quality filtering, which involved removing SNPs with a read depth less than three, genotype quality below 20, and sites with a missingness of more than 20%, resulted in 235,300 SNPs ready for exploratory analyses and further analysis-dependent filtering steps. To ensure similar evolutionary history across all SNPs, we removed the mitochondrial SNPs (10 SNPs), which left a set of 235,290 nuclear SNPs. These SNPs were further pruned to remove sites in linkage disequilibrium ($r^2 > 0.2$). Hence, 107,214 SNPs were available for further analyses.

Global population structure

Overall, we observed three highly divergent genetic lineages that exhibited no evidence of recent admixture, namely, the Atlantic, Indo-Pacific and Pacific lineages (Figure 1). The Atlantic lineage was comprised of all 63 individuals sampled in the Atlantic basin, the Indo-Pacific lineage comprised 38 individuals sampled from the North Pacific and the Indian Ocean site, and the Pacific lineage was comprised of 60 individuals sampled from both the North and South Pacific. The Indo-Pacific and Pacific lineages were sympatric at three sites in the North Pacific, namely KH1110_02, KH1110_05 and KOK1703_03, with an even distribution of 16 animals representing the Indo-Pacific lineage and 16 animals representing the Pacific lineage (Supplementary Material 1.2). Based on the Principal Component Analysis (PCA) plot on the linkage-pruned set of 107,214 SNPs, we observed three genetic clusters with no intermediates between them (Figure 1B). The first two Principal Component (PC) axes comprised 39.7% of the total genetic variation, and PC1 sep-

arated the Atlantic from the Indo-Pacific and Pacific lineages, whereas PC2 separated the Indo-Pacific and Pacific lineages. Congruently, three ancestral populations were the best supported (FIGURE 1C) as K=3 gave the lowest cross-validation error of 0.232.

The same three genetic clusters were also recovered from the nearest neighbour haplotype co-ancestry matrix generated by fineRADstructure (FIGURE 2). The Atlantic lineage showed three further sub-populations corresponding geographically to North, Equatorial and South Atlantic sampling sites (TABLE 1). Within the Indo-Pacific and Pacific lineages, there were also clusters of higher shared co-ancestry, which were geographically localised. These corresponded to individuals from the Indian Ocean within the Indo-Pacific lineage, and individuals from North Pacific sites within the Pacific lineage (FIGURE 2).

The global weighted F_{ST} of the Atlantic lineage (mean = 0.127) was much higher compared to the Indo-Pacific (mean = 0.0321) and Pacific (mean = 0.0429) lineages, and in all three lineages, global weighted F_{ST} was significantly higher than zero ($p < 0.0000$). This indicates significant population substructure across all lineages, with greater population sub-structure within the Atlantic compared to the Indo-Pacific and Pacific lineages (FIGURE 3C).

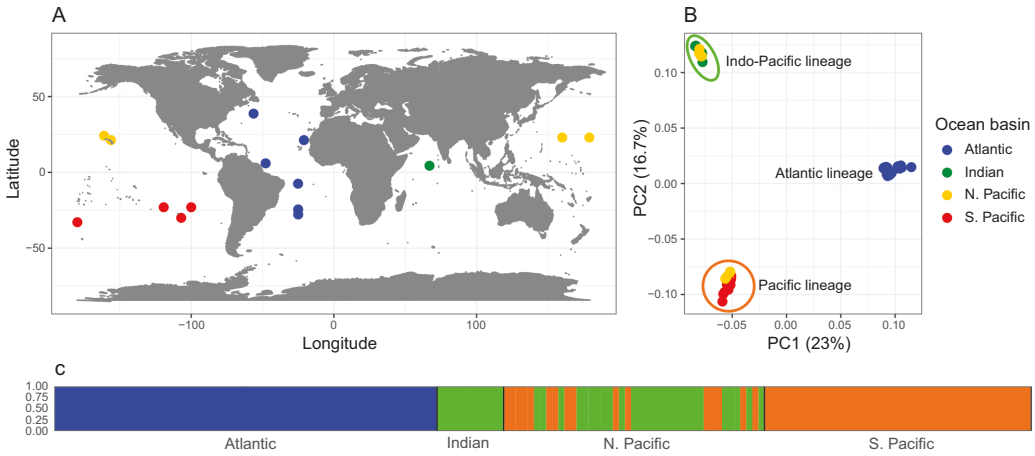


FIGURE 1 (A) Geographic locations of the 14 sampling sites of *Limacina bulimoides*, coloured according to ocean basin: Atlantic, Indian, North (N.) and South (S.) Pacific Ocean (see legend). (B) Principal Component Analysis (PCA) plot based on 107,214 nuclear SNPs from 161 individuals, illustrating the genetic structure at large scale within this species. The three main genetic lineages are indicated with coloured circles: Atlantic (blue), Indo-Pacific (green), Pacific (orange). (C) Admixture plot of *L. bulimoides* populations from the Atlantic, Indian and Pacific Ocean for three putative ancestral populations (K=3) that had the lowest cross-validation error. Each line represents one individual, while colour represents admixture proportion corresponding to the ancestral population (Atlantic lineage in blue, Indo-Pacific lineage in green, and Pacific lineage in orange).

The mitochondrial COI haplotype network (Supplementary Material 4.1.2) showed that the COI sequences are unique across all individuals, with the exception of one shared sequence between two individuals from the Atlantic Ocean. Specimens of the Atlantic lineage could be distinguished readily from the Indo-

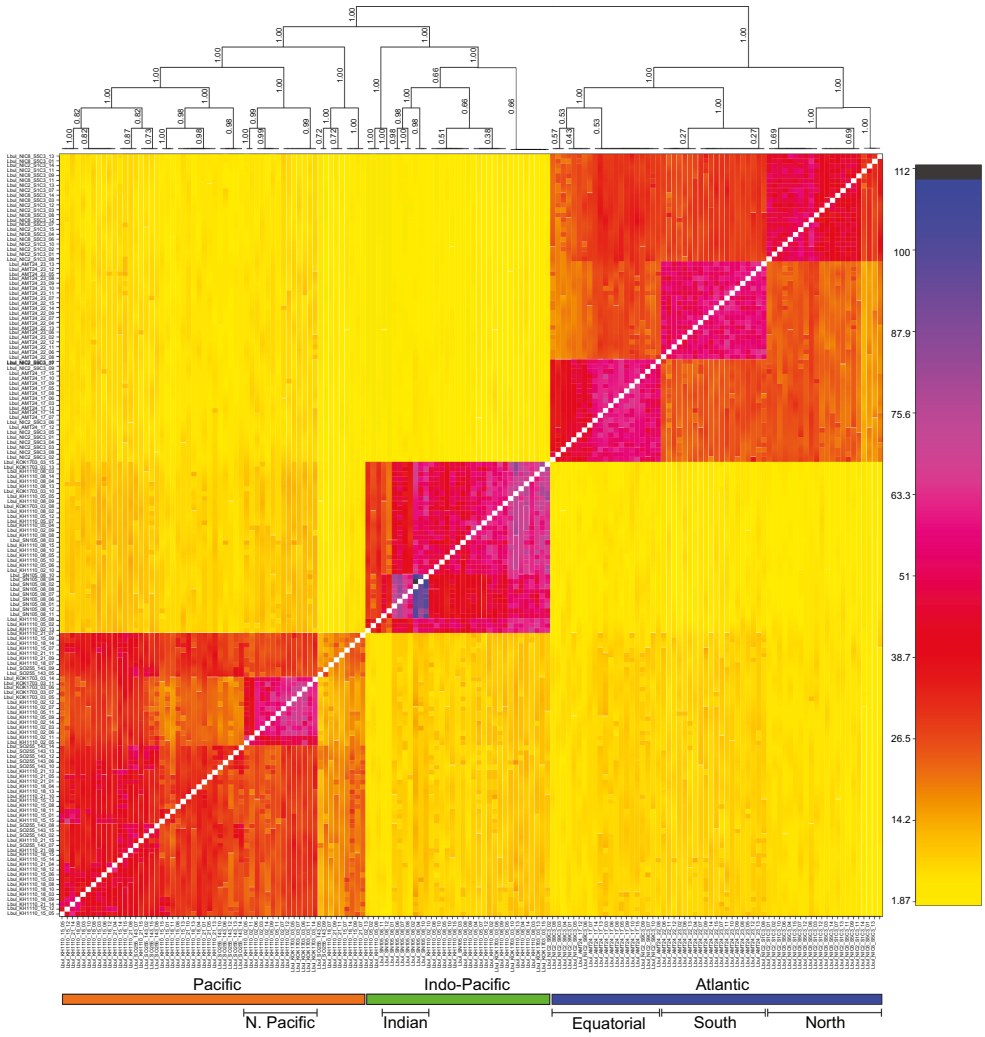


Figure 2 Co-ancestry matrix showing numbers of single-nucleotide polymorphism (SNP) loci with estimated shared co-ancestry for 156 individuals, coloured according to their co-ancestry value (see legend: black/blue colours indicate more relatedness while yellow indicates less relatedness). The three main genetic lineages (Atlantic, Indo-Pacific and Pacific) can be identified, along with finer scale structure within each lineage. The Atlantic lineage (blue) can be further sub-divided into three geographic regions with higher co-ancestry: North, Equatorial and South Atlantic. Within the Indo-Pacific lineage (green), there is higher co-ancestry within the Indian Ocean locality compared to the North Pacific sampling sites. Within the Pacific lineage (orange), the North Pacific individuals have a higher co-ancestry than with the South Pacific lineages.

Pacific and Pacific lineages with at least 62 nucleotide substitutions. Congruent with our previous study (Choo et al. 2021), specimens from the North and Equatorial Atlantic (stations NIC2_S1C3, NIC2_S9C3, NIC8_S5C3 and AMT24_17) could be separated from specimens from the South Atlantic (stations AMT24_22 and AMT24_23) by 17 nucleotide substitutions. However, there are exceptions, namely five individuals that were sampled in the South Atlantic grouped with the North and Equatorial haplotype cluster in the mitochondrial network. Interestingly, these individuals appeared on the outside of the 'wrong' haplotype cluster (Supplementary Material 4.1.2) but grouped according to their geographic sampling location using the nuclear SNP dataset (Figure 2). This indicates that these individuals are not recent migrants or expatriates, transported across dispersal barriers in the Atlantic Ocean, but could represent examples of introgression of mitochondrial DNA. The Indo-Pacific and Pacific lineages could not be clearly distinguished in the mitochondrial COI network. We observe a central cluster of individuals belonging to the Pacific lineage, composed mostly of individuals sampled from the South Pacific. From this central cluster, there are different haplotype groups consisting of individuals belonging to either the Pacific or the Indo-Pacific lineage.

Genetic diversity

Individual heterozygosity varied across the three genetic lineages ($F_{(2,158)}$ = 258.7, $p < 2.00e^{-16}$), with the lowest mean heterozygosity of 0.0238 for the Atlantic lineage (N = 63), followed by 0.0240 for the Indo-Pacific lineage (N = 38), and the highest mean heterozygosity of 0.0311 in the Pacific lineage (N = 60). The Pacific lineage had significantly higher heterozygosity compared to the other two lineages (Tukey HSD: Atlantic – Indo-Pacific: p = 0.915, Atlantic – Pacific: p < 0.0000, Indo-Pacific – Pacific: p < 0.0000) (Figure 3A). The Indo-Pacific lineage had the lowest nucleotide diversity (mean = 0.00492), followed by the Atlantic lineage (mean = 0.00529) and the Pacific lineage (mean = 0.00658). Nucleotide diversity distributions differed significantly between the lineages ($F_{(2,8353)}$ = 145.2, $p < 2.00e^{-16}$, Tukey HSD: Atlantic – Indo-Pacific: p = 0.000945, Atlantic – Pacific: p < 0.0000, Indo-Pacific – Pacific: p < 0.0000) (Figure 3B).

Divergence-time and population size

The SNAPP tree including three *L. bulimoides* genetic lineages and the *L. trochiformis* outgroup supported a split of the ancestral *L. bulimoides* population into the Atlantic and the Pacific/Indo-Pacific lineages (Figure 4). The total number of biallelic loci accepted by SNAPP for the analysis was 1,279, instead of the whole dataset containing 19,669 loci, because SNAPP only accepts alleles that are shared across all four taxa, including the outgroup. The tree supported the monophyly of the Atlantic lineage and the Indo-Pacific/Pacific lineage, with their divergence estimated at 1.20 million years ago (Mya) (95% Highest Posterior Density (HPD) = 0.737- 1.76 My). The Indo-Pacific and Pacific lineages split from one another with a mean age of 0.978 My (95% HPD = 0.603-1.43 My). The average estimated effective population size (N_e)

was the highest for the Pacific lineage (N_e = 6.36 x10^6, 95% HPD = 4.17-8.56 x10^6), followed by the Atlantic lineage (N_e = 3.02 x10^6, 95% HPD = 2.12-3.86 x10^6) and was lowest for the Indo-Pacific lineage (N_e = 1.61 x10^6, 95% HPD = 1.23-2.02 x10^6).

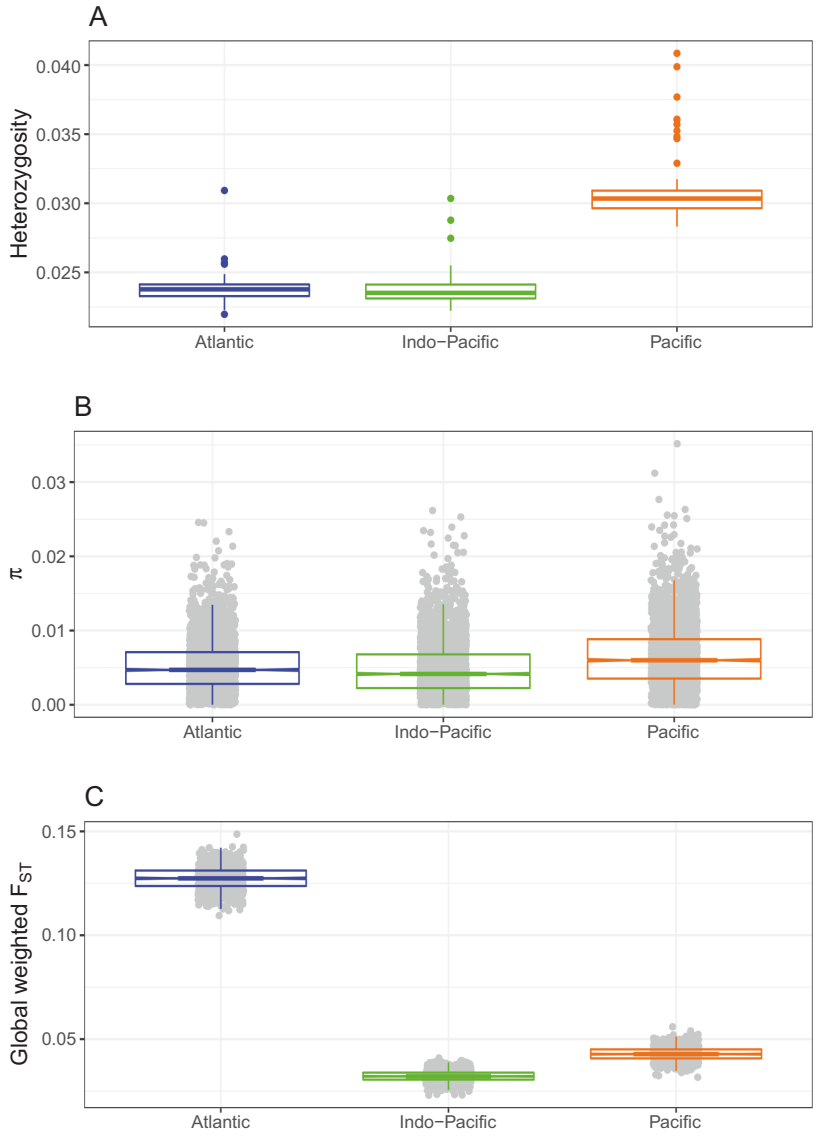


FIGURE 3 Distribution of heterozygosity (A), nucleotide diversity (π) (B) and global weighted F_{ST} (C) for each of the three genetic lineages of *Limacina bulimoides*: Atlantic (N= 63), Indo-Pacific (N=38) and Pacific (N=60). The lower and upper hinges of the boxplots represent the first and third quartile and span in the inter-quartile range (IQR), while median is indicated by the line in between the hinges of the boxplot. The whiskers of the plot extend to values no further than 1.5*IQR.

We found that all three lineages had roughly similar demographic histories using stairway plot analyses (FIGURE 5). Up until 200,000 years ago, there was a steep increase in effective population size (N_e) to ~8 million, which roughly coincided with Marine Isotope Stage (MIS) 6, a period with global glacial conditions. This was followed by subsequent stabilisation in N_e during MIS 5, which had warmer conditions, and MIS 2-4, which is also known as the Last Glacial Period. From MIS 1, the start of the Holocene interglacial at 10,000 years ago, N_e across the three lineages decreased. The N_e of the Indo-Pacific lineage remained stable in population size from 3,000 years ago to the present (N_e = 1.28 x10^6, 95% confidence interval (CI) = 2.56-4.91 x10^6), whereas the N_e of the Atlantic and Pacific lineages continued to decrease to the present (Atlantic: N_e = 5.61 x10^4, 95% CI = 2.34 x10^4-1.99 x10^5; Pacific: N_e = 5.62 x10^4, 95% CI = 1.38 x10^4-1.31 x10^5).

MORPHOLOGICAL VARIATION

Based on the repeatability analysis with 30 individuals landmarked independently by two observers, the relative warps (RWs) 1, 2, 3, 5, 10 and 11 were selected as repeatable parameters to be used in the geometric morphometric analyses of shell

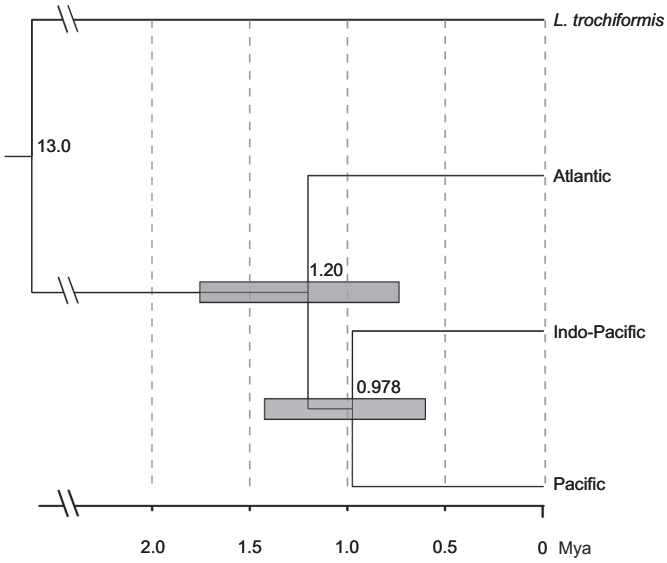


FIGURE 4 Maximum clade credibility tree of the SNAPP phylogeny traces divergence age of three genetic lineages of *Limacina bulimoides* to ~1 Mya. The phylogeny of the Atlantic, Indo-Pacific and Pacific *L. bulimoides* and *L. trochiformis* was based on 1,279 thinned, biallelic nuclear SNPs. Two individuals of *L. trochiformis* and ten individuals per lineage of *L. bulimoides* were included in this tree. The long branches at the root were truncated to allow for better visualisation of the recent divergences. The mean age of each node is labelled above the node, and the bars indicate the 95% highest posterior densities. The scale at the bottom represents time in million years ago (Mya).

shape. Of the 161 *L. bulimoides* individuals, 159 individuals had the complete set of 11 (semi-)landmarks and were included in the analysis. The six repeatable RWs explained 83.75% of shell shape variation (Supplementary Material 4.2.1). Most of the geometric morphometric variation was due to changes in shell width as shown in RW1 accounting for 51.05% of the total variation, and relative size of the aperture as shown in RW2 explaining 18.16% of the total variation (Figure 6A). Though the shells shapes of the three genetic lineages overlapped to a large extent, the distributions were significantly different (PERMANOVA: $F_{(2,156)} = 18.99$, $R^2 = 0.196$, $p = 1e\text{-}04$, Table 2). Likewise, shell shape was significantly different between the three genetic lineages in a canonical variate analysis (One-way ANOVA: $F_{(2,156)} = 41.16$, $p = 4.42e\text{-}15$) (Supplementary Material 4.2.2, Table 2). Grouping specimens according to their sampling location also resulted in significant differences in shell shape (Supplementary Material 4.2.3 and 4.2.4), but this is to be expected because sampling locations are strongly correlated with genetic lineages.

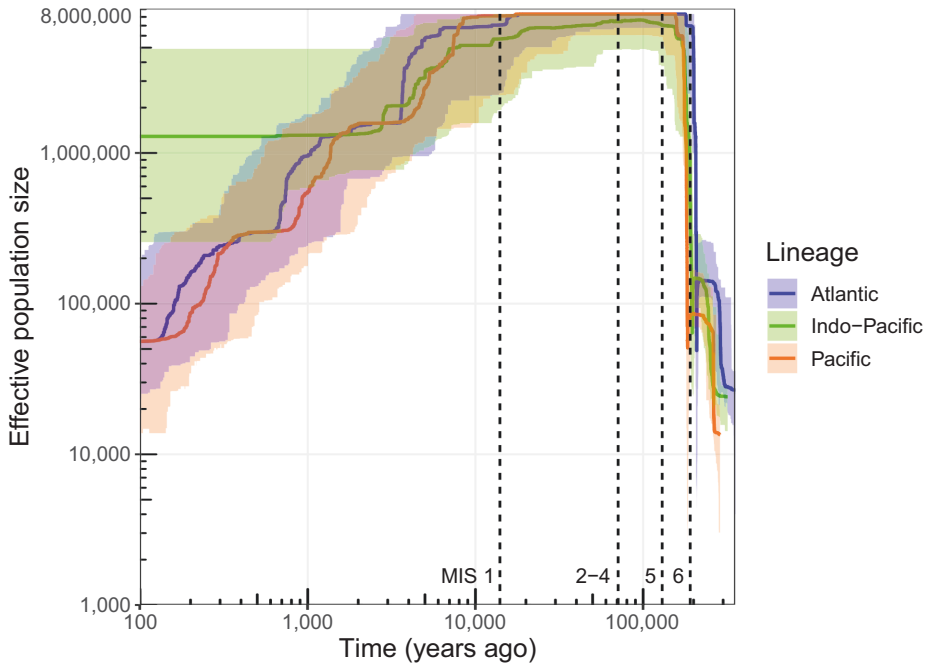


Figure 5 Effective population size (N_e) of the three *Limacina bulimoides* lineages through time as reconstructed by a stairway plot. The x-axis is the number of years before present, assuming a generation time of one year. Estimates of N_e are coloured per lineage, with the line in bold showing the median and the lighter coloured area representing the 95% confidence intervals. Axes are log-scaled for clearer visualisation of recent histories. Plot was truncated from 0 to 100 years ago as sample sizes are not sufficient for reconstructing very recent events as in (Liu and Fu, 2020). Dotted lines indicate the temporal boundaries between Marine Isotope Stages 1 (Holocene), 2-4 (the Last Glacial Period), 5 and 6.

TABLE 2 Pairwise PERMANOVA and Tukey HSD of canonical variate results for comparisons of shell shape variation between genetic lineages of *Limacina bulimoides* (Atlantic, Indo-Pacific and Pacific). All pairwise comparisons are significant, after strict Bonferroni corrections (α = 0.05, p < 0.0167). Significant p-values are indicated in bold.

	PERMANOVA				Tukey HSD			
Pairwise comparison	df	F	R^2	p	Difference	Lower	Upper	p
Atlantic - Indo-Pacific	(1,97)	13.1	0.119	**6.00e-04**	1.35	0.647	2.05	**3.11e-06**
Atlantic - Pacific	(1,121)	6.49	0.0509	**0.0139**	2.30	1.70	2.91	**0**
Indo-Pacific - Pacific	(1,94)	59.6	0.388	**1.00e-04**	0.959	0.253	1.66	**4.48e-04**

We observed large variability in shell colour within our samples. The *L. bulimoides* specimens ranged in colour from almost completely white to beige to reddish-brown (FIGURE 6B, SUPPLEMENTARY MATERIAL 4.3.1). In most coloured individuals, pale beige or red pigmentation was found on the inner aperture, lower half of the aperture, and sometimes on the sutures of the shell. Tissue pigmentation also varied widely, with dark grey tissue that is visible through the somewhat transparent shell observed in specimens from the North Atlantic and Indian Ocean, while almost completely white tissue (and shell) was mostly recorded from the South Pacific (SUPPLEMENTARY MATERIAL 4.3.1). Since the colours of the shell and tissue are confounded with each other and difficult to distinguish, we limited the statistical analyses to shell aperture colour, which is not affected by tissue pigmentation.

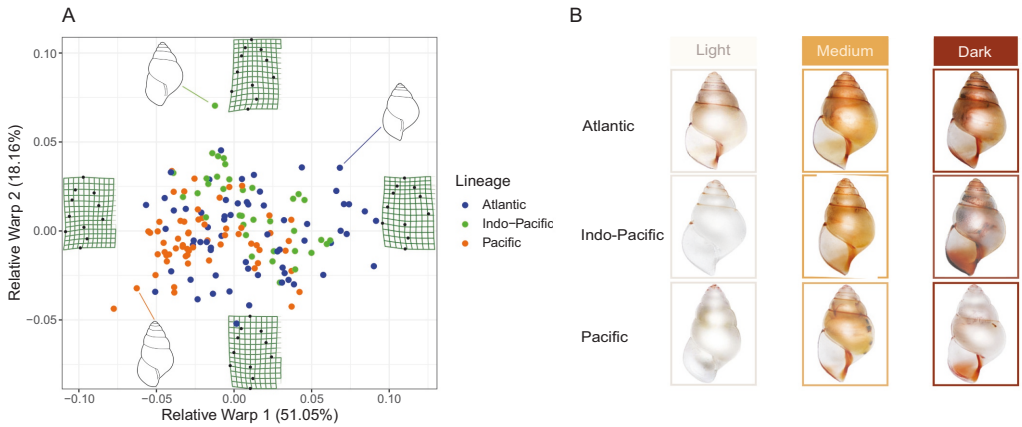


FIGURE 6 Variation in shell shape and aperture colour across the three lineages of *Limacina bulimoides*. (A) Shell shape variation of 159 *L. bulimoides* specimens, categorised into the three genetic lineages (Atlantic, Indo-Pacific and Pacific). Shell shape variation is visualised with relative warp axes 1 and 2, explaining 51.05% and 18.16% of total shell shape variation, respectively. Extremes of both relative warp axes are shown with the thin plate spline images, with each black dot corresponding to a shell landmark. Line drawings of example individuals from each genetic lineage are shown: Atlantic (Lbul_NIC2_S9C3_04), Indo-Pacific (Lbul_SN105_08_02), and Pacific (Lbul_KH1110_18_03). (B) Illustration of the intensity of aperture colour observed, arranged from light, medium to dark apertures in each of the three lineages (see also SUPPLEMENTARY MATERIAL 4.3.1).

Aperture colour was not randomly distributed with respect to genetic lineage (p = 0.0005, two-sided), nor sampling location (p = 0.0005, two-sided) (Supplementary Material 4.3.1). Atlantic specimens had mainly pink and red-brown apertures while North Pacific specimens were highly pigmented with mainly tan and red-brown apertures. Specimens from the South Pacific had less pigmented apertures, with either transparent or pink apertures, and Indian Ocean specimens had highly pigmented shells with red-brown apertures. Aperture colour was often consistent within sampling station (six of the 15 stations) but could also vary (e.g., KH1110_05 from the North Pacific had all four aperture colours represented (Supplementary Material 4.3.2).

We closely examined morphological variation in the individuals of the two lineages that co-occur in the North Pacific, namely the Indo-Pacific lineage (n = 16) and Pacific lineage (n = 16) (Figure 7). We found that there were no significant differences in shell shape between the two lineages despite their distinct genetic

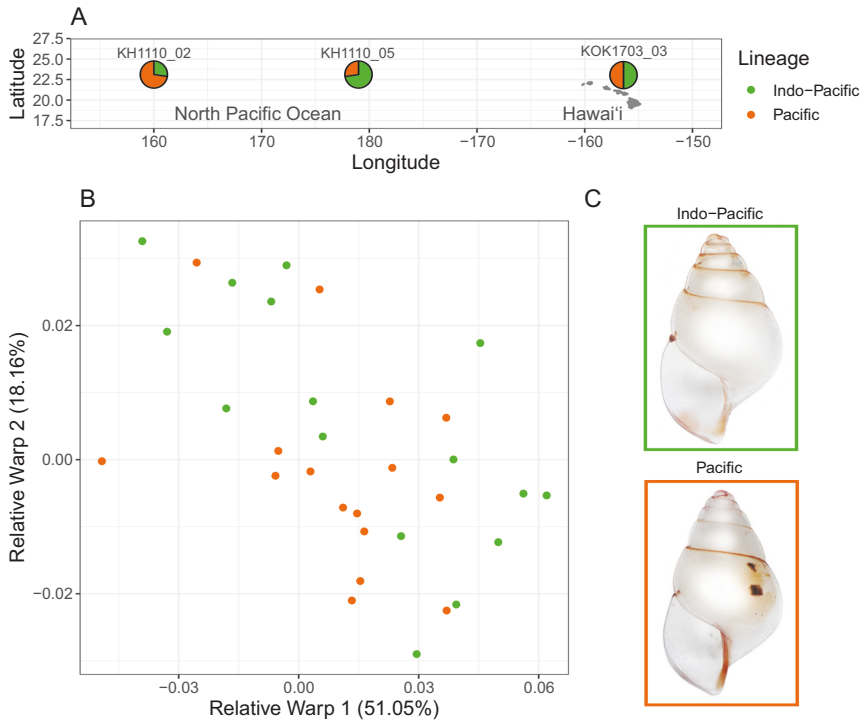


Figure 7 Distribution (A) and morphological variation (B, C) of the sympatric genetic lineages of *Limacina bulimoides* in the North Pacific with 16 individuals from the Indo-Pacific lineage (green) and 16 individuals from the Pacific lineage (orange). (B) Shell shape variation as visualised with relative warps 1 and 2, which explain 51.05% and 18.16% of total shell shape variation, respectively. (C) Photographs of an example individual from each genetic lineage are shown, with dark pigmented spots on the tissue, visible through the transparent shell, of the specimen from the Pacific lineage (see also Supplementary Material 4.3.1 and 4.3.3).

backgrounds (Figure 7B; PERMANOVA: $F_{(1,30)} = 3.75$, $R^2 = 0.111$, $p = 0.0612$). From visual observations of the individual photographs, we noticed that the Indo-Pacific and Pacific lineages could potentially be differentiated by the presence of dark pigmented spots on the tissue of Pacific lineage individuals only, which is visible through their transparent shell (Figure 7C). Dissection of additional individuals from the same samples showed that these pigmented spots were localised on the margin of their 'wing-feet' or parapodia (Supplementary Material 4.3.3). These pigmented spots were not observed in the photographs of other individuals belonging to the Pacific lineage but sampled in a different location, in the South Pacific (Supplementary Material 4.3.1).

Discussion

The shelled pteropod *Limacina bulimoides* is composed of three main genetic lineages (Atlantic, Indo-Pacific and Pacific) that could be considered as distinct biological species since they appear reproductively isolated with no evidence of recent gene flow or hybridisation. These three lineages diverged approximately one million years ago (Mya) during the mid-Pleistocene, and we observed further geographical structuring within each of the three lineages.

While there are subtle morphological differences between the three main lineages, we are unable to disentangle the relative effects of genetics, environment and ecology on shell morphological characters in *L. bulimoides*. The three lineages cannot be separated on the basis of morphological traits such as shell shape or colour due to the overlapping, albeit significantly different, shell shape distributions and highly variable shell colour across sampling locations (Figure 6, Supplementary Material 4.3.2). Shell shape in molluscs is unlikely to be solely affected by genetic variation, especially in taxa that have high dispersal potential and occupy heterogeneous habitats. Instead, shell shape may have a phenotypically plastic component linked with life history and environmental conditions (Hoffman et al., 2010; Hollander et al., 2006; Mariani et al., 2012; Zieritz et al., 2010). Pteropod shell shape is an important trait that directly affects their sinking and swimming speeds, manoeuvrability, and their resulting ability to navigate the water column for food and evade predators (Karakas et al., 2020). In other molluscs, shell shape has been shown to be correlated with environment, either as genetically inherited differences or phenotypically plastic traits. For instance, in various species of intertidal snails with a 'crab' or 'wave' ecotype, e.g., *Littorina* and *Nucella* (Guerra-Varela et al., 2009; Hollander and Butlin, 2010; Johannesson, 2003; Rolán et al., 2004), or in *Mytilus* which exhibits shell shape plasticity as a response to environmental parameters like temperature and food (Telesca et al., 2018).

The non-random variation in shell colour across sampling locations (Supplementary Material 4.3.1) may be due to pigments incorporated from their diet, which is composed of phytoplankton, microbes and particulate matter trapped with their mucous web (Conley et al., 2018). *Limacina* can feed selectively by mov-

ing the cilia on their wings and mantle lining to sort and reject food particles (Lalli and Gilmer, 1989) and can control their vertical distribution through diel vertical migration. These differences in food choice and vertical habitat may lead to variable shell colour among individuals from the same sampling station. Production of shell colour has been suggested to be energetically costly across molluscs (Williams, 2017), but pteropods, with their unique planktonic life style, may be subject to other selective trade-offs such as being transparent to remain inconspicuous to visual predators in the water column (Johnsen, 2001) or possessing red pigment for protection against UV radiation (Hansson, 2000).

Since the Indo-Pacific and Pacific lineages are sympatric in the North Pacific, and yet have remained genetically distinct with no sign of recent gene flow, they are probably reproductively isolated. The possible modes of reproductive isolation in *L. bulimoides* include prezygotic, such as habitat, temporal, and behavioural isolation, and postzygotic means, such as gamete isolation and non-viability of hybrids. Specimens from these three stations (KH110_02, KH1110_05, and KOK1703_03) were collected from oblique tows in surface waters (maximum depth was 370 meters), therefore it is unclear if the two lineages may have distinct depth habitats and/or corresponding diet preferences, and thus could experience habitat or ecological isolation. *Limacina bulimoides* have internal fertilisation and it is likely that they have evolved mate recognition mechanisms in order to locate each other in the pelagic environment. While they are cryptic in terms of their shell shape, the two sympatric lineages in the North Pacific could potentially be visually differentiated by the presence of tissue pigmentation on their parapodia (Supplementary Material 4.3.3). These 'wing' spots seem analogous to wing pigmentation found on butterflies and flies, which could be a distinguishing trait that mediates species recognition (Gompel et al., 2005; Wiernasz and Kingsolver, 1992). However, it is questionable whether *Limacina*, which are also commonly referred to as 'sea-butterflies', can actually see such spots, and the extent to which they rely on their sight for possible mate recognition. *Limacina* possess eyes and optic nerves (Laibl et al., 2019), and are capable of detecting changes in light levels to trigger diel vertical migration (Cohen and Forward Jr, 2016) but there is a lack of information on the types of cues used by pteropods to recognise conspecifics. In other planktonic species, such as copepods, pheromone trails and swimming patterns are used for mate finding (Bagøien and Kiørboe, 2005; Kiørboe, 2007), which may also serve to mediate the reproductive isolation between sibling species, as seen in other marine organisms like isopods, stomatopods and amphipods (Palumbi, 1994).

There is insufficient evidence to conclude whether the Pacific and Indo-Pacific lineages evolved in sympatry or have arrived at their present-day distribution through secondary contact. We do not see a signal of recent range expansion and increased N_e in either of the two lineages, which would suggest stable population sizes and divergence in sympatry, although additional sampling locations should be included for more reliable demographic inferences. The absence of the Pacific lin-

eage from the Indian Ocean, despite potential connectivity with the North Pacific, where both lineages are present, could be the result of either incomplete sampling, niche incumbency effects, where a new species is prevented from expanding into a suitable habitat due to competitive exclusion by an existing species with similar niche (e.g., Weiner et al., 2014), or recent extinction in the Indian Ocean, which has been recorded in the fossil record of another pteropod species (Wall-Palmer et al., 2014, 2016). More detailed sampling, in the Indian and Pacific Oceans, will be needed to resolve the geographical boundaries and extent of overlap between the two lineages and understand the causes behind their modern-day distribution. Sampling of *L. bulimoides* with depth stratified collection techniques, metabarcoding of microbiome and gut contents, examination of their radula, transcriptome sequencing and observations of their reproductive timings will be also be needed to gain more insight into the possible modes of current reproductive isolation between the two lineages.

We note the usefulness of genome-wide data, as compared to mitochondrial DNA or barcoding genes, in detecting population structure and assessing species boundaries in non-model organisms. With the data from genome-wide SNPs, we can conclude that there has not been recent hybridisation or introgression between the three main lineages, which points towards them being reproductively isolated. If we were to solely rely on the mitochondrial COI barcoding region in *L. bulimoides* as an indicator of overall genomic divergence, we would have underestimated the levels of genomic divergence between the Indo-Pacific and Pacific lineages and the substructure within the Atlantic lineage, even though it is, in fact, a highly diverse marker (8.69% nucleotide diversity). In other studies using barcode regions, species delimitations have been made based on taxon-specific thresholds of levels of divergence beyond which reproductive isolation is expected (Krug et al., 2013; Lefébure et al., 2006; Young et al., 2017). While this method is highly effective for detecting species in broad-scale barcoding studies without prior information, the COI gene may not evolve at a constant rate over time across all taxa (Thomas et al., 2006). In addition, using only one or a few barcoding genes to delimit species based on their divergence distances may not indicate reproductive isolation or lead to further understanding on the process of speciation (Freeman and Pennell, 2021). Speciation with gene flow can occur via selection on many unlinked genes, leading to genomic islands and barriers to gene flow (Feder and Nosil, 2010), and can be detected with genome-wide analyses. With reductions in costs for sequencing and increased availability of genetic resources, whole genome data can be used to provide additional information to identify trends and changes in genome architecture, a key factor facilitating rapid adaptive phenotypic shifts (Koch et al., 2021; Therkildsen et al., 2019). The additional genomic and recombination rate information can also be used to identify if the traits under selection have a polygenic architecture, which is associated with the maintenance of reproductive isolation in sympatry and promotion of speciation in the presence gene flow (Kautt et al., 2020; Martin et al., 2019).

While we have only focused on SNPs in this study, further analyses of structural variants across the genome can allow us to identify genomic architecture that may play a role in reproductive isolation between the lineages. For this to be possible, we would require a more contiguous reference genome for *L. bulimoides* as compared to what is currently available (Choo et al., 2020).

At the global scale, the spatial genetic structuring of the *L. bulimoides* lineages is congruent with an allopatric model of speciation, with the timing of lineage divergence being roughly concordant with glacial events. The divergence of *L. bulimoides* lineages at around 1 Mya (Figure 4) coincides with the timing of divergence in a mesopelagic copepod species (Andrews et al., 2014) during the mid-Pleistocene transition (0.6-1.2 Mya). This period was characterised by global cooling, lengthening of glacial stages from 41k year cycles to 100k year cycles, changing ocean circulation and productivity, and the evolution of many terrestrial and marine biota (Clark et al., 2006; Elderfield et al., 2012; Kender et al., 2016; McClymont et al., 2013). Changes in ocean circulation during the mid-Pleistocene transition could have facilitated the physical separation and subsequent divergence of the three lineages across the various ocean basins. These changes include the reduced exchange between the Indian and Atlantic Oceans via the Agulhas leakage due to the northward migration of the Subtropical Front towards the Agulhas Plateau (Caley et al., 2012; Cartagena-Sierra et al., 2021), and the connection between the Pacific and Indian Ocean due to the weakening of the Indonesian Throughflow (Petrick et al., 2019). Similar geographical structuring across ocean basins for other circumglobal warm-water plankton species has been attributed to both physical (e.g., ocean currents or continental landmasses) and ecological (species-specific interactions with oceanographic gradients) barriers (e.g., Burridge et al., 2019; Goetze et al., 2015; Hirai et al., 2015), although it is unknown if these structured populations arose at the same time. Within the Pacific and Atlantic basins, habitat discontinuities like the mesotrophic equatorial upwelling waters, have been described as ecological dispersal barriers for subtropical copepods (Goetze, 2005; Goetze et al., 2017) and the pteropod genus *Cuvierina* (Burridge et al., 2015). Further genome-wide studies in other circumglobal species should be conducted to identify if there are congruent drivers of historical divergence among holoplanktonic organisms.

The *L. bulimoides* species complex is likely to be resilient to ocean changes, based on their high genomic variability and the absence of population bottlenecks throughout their evolutionary history across multiple Pleistocene glacial-interglacial transitions. Stable population sizes in *L. bulimoides* were associated with the Last Glacial Period (MIS2-4), as well as MIS5. While the MIS 5 includes the Eemian / Last Interglacial Period (129-116k years), which was characterised by global temperatures that were warmer than the present day, there was little to no increase in sea surface temperatures (SST) in the tropics based on reconstructions of sea surface temperatures derived from faunal and floral assemblages, foramineral and alkenone proxies (Turney et al., 2020). Stable abundances of *L. bulimoides* were also

recorded with sediment core data in the Caribbean Sea, which had low variation in SST across both glacial and interglacial periods (Wall-Palmer et al., 2014). At the beginning of the Holocene, between 19-7 k years ago (boundary of MIS1/2), temperatures increased and global sea levels rose rapidly (Lambeck and Chappell, 2001), which was associated with a decrease in N_e. In the stairway plot, N_e is calculated as a function of coalescent time (Liu and Fu, 2020), and a decrease in time to coalescence can indicate either a decrease in population size or a decrease in population connectivity due to population divergence (Foote et al., 2016; Mazet et al., 2016). It is possible that the increase of temperature and sea-levels have facilitated further population divergence to result in the geographically separated population substructure we observe in the present day. A major caveat is that the timings of divergence and demographic changes are a function of generation time, which was set to one year, and may vary if a more accurate value of generation time is obtained for this species. Since the stairway plot reconstructions for the Indo-Pacific lineage used fewer haploid sequences (40 sequences) than the Atlantic (56 sequences) and Pacific (64 sequences) lineages, there is less resolution in reconstructing more recent demographic events for the Indo-Pacific lineage, which may explain its difference in trend in N_e compared to the other two lineages. Given that most of our targeted regions are coding regions in the genome, we expect that these demography inferences may be affected by background selection (Ewing and Jensen, 2016) leading to a lower than expected recent N_e for the lineages as compared to estimates from the conceptually different SNAPP analysis, because of purifying selection against mutations at linked sites leading to a loss of diversity. The relative N_e estimates from the stairway plot are also sensitive to population substructure, which is lower within the Indo-Pacific lineage compared to the Atlantic and Pacific lineages (Figure 3C). This can result in an inflation of Indo-Pacific N_e based on the stairway plot whereas the SNAPP N_e estimates showed that the Indo-Pacific had the lowest effective population size among all three lineages.

Despite the potential for global dispersal in *L. bulimoides*, we observed diversification of lineages across oceanic basins that likely originated from the mid-Pleistocene transition. These lineages are probably distinct species with strong reproductive isolation between them, and have survived through periods of glacial-interglacial transitions representing a wide range of oceanographic conditions. There are slight differences in shell shape between the lineages, but shell shape alone cannot be used as a taxonomic character, and it is unclear whether environmental or genetic factors have a greater impact on shell morphology. Even though their effective population sizes have decreased since the start of the Holocene, the three lineages still possess high levels of standing genetic variation and nucleotide diversity, which should be useful for adaptation to future changes (Bernatchez, 2016; Bitter et al., 2019; Schluter and Conte, 2009). However, it is unclear whether pteropods and other planktonic calcifiers can cope with the rate of ongoing ocean changes, including anthropogenic carbon emission rates that are unprecedented

since at least 66 million years, leading to increasing ocean acidification (Zeebe et al., 2016). Further genome-wide analyses of *L. bulimoides* in the currently under-sampled Indian and Pacific oceans should be conducted to improve understanding into the drivers of divergence in this holoplanktonic species and estimate their capacity to adapt to future conditions.

Acknowledgements

We thank L. Mekkes, A. Burridge and M. Jungbluth for assistance at sea, and all captains and crews of the ocean expeditions for their support and assistance. We thank D. Wall-Palmer for providing us Indian Ocean and South Pacific samples from the SN105 and SO255 cruises and A. Tsuda for providing the Pacific samples from the 2011/2012 R/V Hakuho-Maru KH-11-10 cruise. We warmly acknowledge M. Malinsky, M. Matschiner and E. Trucchi for their advice on data analyses, W. Renema and D. Wall-Palmer for helpful discussions regarding the manuscript, and M. Kopp for guidance in the lab. This research was supported by the Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to K.T.C.A.P. The Netherlands Initiative Changing Oceans (NICO) expedition on *RV Pelagia* was also funded by NWO and the Royal Netherlands Institute for Sea Research (NIOZ). Further fieldwork was supported by NSF grants OCE-1029478 and OCE-1338959 to E.G. The *R/V Sonne* cruise SO255 was funded by the German Federal Ministry of Education and Research (BMBF; grant 03G0255A), and the SN105 cruise on board the ORV Sagar Nidhi was funded by the Indian National Centre for Ocean Information Services (INCOIS), Ministry of Earth Sciences, India, as the first cruise of the second International Indian Ocean Expedition (IIOE-2). The Atlantic Meridional Transect is funded by the UK Natural Environment Research Council through its National Capability Long-term Single Centre Science Programme, Climate Linked Atlantic Sector Science (grant number NE/R015953/1). This study contributes to the international IMBeR project and is contribution number 368 of the AMT programme.

Supplementary material

Contents

1. Specimens and genetic lineage assignment

1.1. Sampling overview information

Table S1 Sampling location, station name, museum voucher code, short read archive (SRA) and mitochondrial cytochrome oxidase I (COI) accessions, capture efficiency statistics, assigned lineage and inclusion in analyses for each individual.

Specimen ID	Location	Station	Museum voucher code	SRA	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth	Lineage	SNAPP/ stairway plot	Morpho-metrics
Lbul_AMT24_17_03	Atlantic	AMT24_17	RMNH.MOL.341275	SAMN20293115	MZ542566	8548346	2641394	30.9	86.5	162.13	Atlantic	No	Yes
Lbul_AMT24_17_05	Atlantic	AMT24_17	RMNH.MOL.340317	SAMN20293116	MZ542567	8619139	4035590	46.8	91.7	267.48	Atlantic	No	Yes
Lbul_AMT24_17_06	Atlantic	AMT24_17	RMNH.MOL.340318	SAMN20293117	MZ542568	9525743	5112712	53.7	93.6	348.92	Atlantic	No	Yes
Lbul_AMT24_17_07	Atlantic	AMT24_17	RMNH.MOL.340319	SAMN20293118	MZ542569	5945465	3128096	52.6	93	211.7	Atlantic	No	Yes
Lbul_AMT24_17_08	Atlantic	AMT24_17	RMNH.MOL.340320	SAMN20293119	MZ542570	13129266	7287498	55.5	88.8	467.92	Atlantic	Yes	Yes
Lbul_AMT24_17_09	Atlantic	AMT24_17	RMNH.MOL.340321	SAMN20293120	MZ542571	10888222	5872132	53.9	87.9	371.76	Atlantic	Yes	Yes
Lbul_AMT24_17_10	Atlantic	AMT24_17	RMNH.MOL.340322	SAMN20293121	MZ542572	13675828	7318800	53.5	87.2	458.84	Atlantic	Yes	Yes
Lbul_AMT24_17_12	Atlantic	AMT24_17	RMNH.MOL.340323	SAMN20293122	MZ542573	4963728	2581066	52.0	94	177.65	Atlantic	No	Yes
Lbul_AMT24_17_13	Atlantic	AMT24_17	RMNH.MOL.340324	SAMN20293123	MZ542574	10474291	5016032	47.9	91.3	331.99	Atlantic	Yes	Yes
Lbul_AMT24_17_14	Atlantic	AMT24_17	RMNH.MOL.340325	SAMN20293124	MZ542575	9865549	5068552	51.4	92	338.06	Atlantic	No	Yes
Lbul_AMT24_17_15	Atlantic	AMT24_17	RMNH.MOL.340326	SAMN20293125	MZ542576	8087831	3514198	43.5	91.3	231.62	Atlantic	No	Yes
Lbul_AMT24_22_04	Atlantic	AMT24_22	RMNH.MOL.341200	SAMN11131477	MZ542577	16895753	4670866	27.6	85.6	284.34	Atlantic	No	Yes
Lbul_AMT24_22_06	Atlantic	AMT24_22	RMNH.MOL.341201	SAMN11131478	MZ542578	6107748	1909732	31.3	82.6	111.5	Atlantic	No	Yes
Lbul_AMT24_22_07	Atlantic	AMT24_22	RMNH.MOL.340345	SAMN20293126	MZ542579	12151985	6428490	52.9	92.4	432.12	Atlantic	No	Yes
Lbul_AMT24_22_08	Atlantic	AMT24_22	RMNH.MOL.341202	SAMN11131479	MZ542580	8193775	3492314	42.6	92.5	234.08	Atlantic	No	Yes
Lbul_AMT24_22_09	Atlantic	AMT24_22	RMNH.MOL.340346	SAMN20293127	MZ542581	7545267	4085630	54.1	94	280.28	Atlantic	No	Yes
Lbul_AMT24_22_11	Atlantic	AMT24_22	RMNH.MOL.340348	SAMN20293128	MZ542582	12123113	6301860	52.0	87.3	398.85	Atlantic	No	Yes
Lbul_AMT24_22_12	Atlantic	AMT24_22	RMNH.MOL.340349	SAMN20293129	MZ542583	16218485	8033820	49.5	87.9	512.47	Atlantic	No	Yes
Lbul_AMT24_22_13	Atlantic	AMT24_22	RMNH.MOL.340350	SAMN20293130	MZ542584	12379217	6530426	52.8	89.1	422.73	Atlantic	No	Yes
Lbul_AMT24_22_14	Atlantic	AMT24_22	RMNH.MOL.340351	SAMN20293131	MZ542585	9636983	4995398	51.8	92.4	336.15	Atlantic	No	Yes
Lbul_AMT24_22_15	Atlantic	AMT24_22	RMNH.MOL.340352	SAMN20293132	MZ542586	10732523	5037788	46.9	91.2	332.12	Atlantic	No	Yes
Lbul_AMT24_23_02	Atlantic	AMT24_23	RMNH.MOL.340353	SAMN20293133	MZ542587	10912937	2885556	26.4	87.7	179.76	Atlantic	No	Yes
Lbul_AMT24_23_05	Atlantic	AMT24_23	RMNH.MOL.340354	SAMN20293134	MZ542588	9060457	4416126	48.7	91.7	292.49	Atlantic	No	Yes
Lbul_AMT24_23_06	Atlantic	AMT24_23	RMNH.MOL.341203	SAMN11131480	MZ542589	10248846	3785256	36.9	83.4	223.16	Atlantic	Yes	Yes
Lbul_AMT24_23_07	Atlantic	AMT24_23	RMNH.MOL.340355	SAMN20293135	MZ542590	12388848	6365196	51.4	92.6	428.94	Atlantic	No	Yes
Lbul_AMT24_23_08	Atlantic	AMT24_23	RMNH.MOL.341204	SAMN11131481	MZ542591	13640477	6294478	46.1	86.4	390.4	Atlantic	Yes	Yes
Lbul_AMT24_23_09	Atlantic	AMT24_23	RMNH.MOL.341205	SAMN11131482	MZ542592	14038163	6982544	49.7	86.7	434.67	Atlantic	Yes	Yes
Lbul_AMT24_23_10	Atlantic	AMT24_23	RMNH.MOL.340356	SAMN20293136	MZ542593	10848376	6083488	56.1	88.8	390.94	Atlantic	No	Yes
Lbul_AMT24_23_11	Atlantic	AMT24_23	RMNH.MOL.340357	SAMN20293137	MZ542594	8918248	4643746	52.1	93.7	317.31	Atlantic	No	Yes
Lbul_AMT24_23_12	Atlantic	AMT24_23	RMNH.MOL.340358	SAMN20293138	MZ542595	15044933	6750610	44.9	90.3	440.47	Atlantic	No	Yes
Lbul_AMT24_23_13	Atlantic	AMT24_23	RMNH.MOL.340359	SAMN20293139	MZ542596	12080446	5990002	49.6	91.3	396.18	Atlantic	No	Yes
Lbul_KH1110_02_03	N. Pacific	KH1110_02	RMNH.MOL.346491	SAMN20293140	MZ542597	8437210	3541496	42.0	96	253.06	Pacific	No	Yes
Lbul_KH1110_02_05	N. Pacific	KH1110_02	RMNH.MOL.346492	SAMN20293141	MZ542598	9175368	3577738	39.0	95.2	253.05	Pacific	No	Yes
Lbul_KH1110_02_06	N. Pacific	KH1110_02	RMNH.MOL.346493	SAMN20293142	MZ542599	10966273	4656112	42.5	96.5	334.46	Pacific	No	Yes
Lbul_KH1110_02_07	N. Pacific	KH1110_02	RMNH.MOL.346494	SAMN20293143	MZ542600	12158794	4919012	40.5	96.1	351.27	Pacific	No	Yes
Lbul_KH1110_02_08	N. Pacific	KH1110_02	RMNH.MOL.346495	SAMN20293144	MZ542601	3228005	1352116	41.9	96.7	97.08	Pacific	No	Yes

TABLE S1 Continued.

Specimen ID	Location	Station	Museum voucher code	SRA	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth	Lineage	SNAPP/ stairway plot	Morpho-metrics
Lbul_KH1110_02_09	N. Pacific	KH1110_02	RMNH.MOL.346496	SAMN20293145	MZ542602	8853744	3565898	40.3	95.8	254.17	Indo-Pacific	No	Yes
Lbul_KH1110_02_10	N. Pacific	KH1110_02	RMNH.MOL.346497	SAMN20293146	MZ542603	9269978	3387162	36.5	94.9	238.67	Indo-Pacific	No	Yes
Lbul_KH1110_02_11	N. Pacific	KH1110_02	RMNH.MOL.346498	SAMN20293147	MZ542604	5963121	2693880	45.2	96.9	194.84	Pacific	No	Yes
Lbul_KH1110_02_12	N. Pacific	KH1110_02	RMNH.MOL.346499	SAMN20293148	MZ542605	11916064	4645316	39.0	95.4	328.41	Pacific	No	Yes
Lbul_KH1110_02_13	N. Pacific	KH1110_02	RMNH.MOL.346500	SAMN20293149	MZ542606	5357671	2340426	43.7	96.7	168.27	Indo-Pacific	No	Yes
Lbul_KH1110_02_14	N. Pacific	KH1110_02	RMNH.MOL.346501	SAMN20293150	MZ542607	8925735	3543636	39.7	96.4	253.04	Pacific	No	Yes
Lbul_KH1110_05_01	N. Pacific	KH1110_05	RMNH.MOL.346502	SAMN20293151	MZ542608	7164357	1458476	20.4	98.3	106.84	Pacific	No	Yes
Lbul_KH1110_05_02	N. Pacific	KH1110_05	RMNH.MOL.346503	SAMN20293152	MZ542609	6791298	2582608	38.0	95.7	183.92	Indo-Pacific	No	Yes
Lbul_KH1110_05_04	N. Pacific	KH1110_05	RMNH.MOL.346504	SAMN20293153	MZ542610	7444714	3114790	41.8	95.6	221.72	Indo-Pacific	No	Yes
Lbul_KH1110_05_05	N. Pacific	KH1110_05	RMNH.MOL.346505	SAMN20293154	MZ542611	8966178	3977166	44.4	94.8	279.15	Indo-Pacific	Yes	Yes
Lbul_KH1110_05_06	N. Pacific	KH1110_05	RMNH.MOL.346506	SAMN20293155	MZ542612	15109965	3109834	20.6	98.4	227.78	Indo-Pacific	No	Yes
Lbul_KH1110_05_07	N. Pacific	KH1110_05	RMNH.MOL.346507	SAMN20293156	MZ542613	11623558	2596938	22.3	97.6	188.54	Indo-Pacific	No	Yes
Lbul_KH1110_05_08	N. Pacific	KH1110_05	RMNH.MOL.346508	SAMN20293157	MZ542614	6823530	2126544	31.2	94.7	148.34	Indo-Pacific	No	Yes
Lbul_KH1110_05_09	N. Pacific	KH1110_05	RMNH.MOL.346509	SAMN20293158	MZ542615	11358903	4379852	38.6	94.7	308.12	Pacific	No	Yes
Lbul_KH1110_05_10	N. Pacific	KH1110_05	RMNH.MOL.346510	SAMN20293159	MZ542616	6907471	3092824	44.8	96.1	221.79	Indo-Pacific	No	Yes
Lbul_KH1110_05_11	N. Pacific	KH1110_05	RMNH.MOL.346511	SAMN20293160	MZ542617	7263229	3315392	45.6	96.8	239.82	Pacific	No	Yes
Lbul_KH1110_05_12	N. Pacific	KH1110_05	RMNH.MOL.346512	SAMN20293161	MZ542618	8034784	3724546	46.4	95.5	264.51	Indo-Pacific	No	Yes
Lbul_KH1110_08_27	N. Pacific	KH1110_08	RMNH.MOL.346513	SAMN20293162	MZ542619	11820493	4703394	39.8	95.2	332.77	Indo-Pacific	Yes	Yes
Lbul_KH1110_08_28	N. Pacific	KH1110_08	RMNH.MOL.346514	SAMN20293163	MZ542620	10004112	4282014	42.8	96.7	308.3	Indo-Pacific	No	Yes
Lbul_KH1110_08_29	N. Pacific	KH1110_08	RMNH.MOL.346515	SAMN20293164	MZ542621	9858669	4433776	45.0	95.6	314.4	Indo-Pacific	Yes	No
Lbul_KH1110_08_30	N. Pacific	KH1110_08	RMNH.MOL.346516	SAMN20293165	MZ542622	8202708	3530060	43.0	96	252.3	Indo-Pacific	No	Yes
Lbul_KH1110_08_33	N. Pacific	KH1110_08	RMNH.MOL.346517	SAMN20293166	MZ542623	11611027	2551628	22.0	98	186.11	Indo-Pacific	No	Yes
Lbul_KH1110_08_34	N. Pacific	KH1110_08	RMNH.MOL.346518	SAMN20293167	MZ542624	7982041	3258516	40.8	96.6	234.14	Indo-Pacific	No	Yes
Lbul_KH1110_08_35	N. Pacific	KH1110_08	RMNH.MOL.346519	SAMN20293168	MZ542625	9603344	3756960	39.1	95.8	267.62	Indo-Pacific	No	Yes
Lbul_KH1110_08_37	N. Pacific	KH1110_08	RMNH.MOL.346520	SAMN20293169	MZ542626	7840905	1750098	22.3	97.7	127.39	Indo-Pacific	No	Yes
Lbul_KH1110_08_38	N. Pacific	KH1110_08	RMNH.MOL.346521	SAMN20293170	MZ542627	10492323	4367256	41.6	96.4	313.35	Indo-Pacific	No	Yes
Lbul_KH1110_08_39	N. Pacific	KH1110_08	RMNH.MOL.346522	SAMN20293171	MZ542628	11879763	4102738	34.5	95.2	289.18	Indo-Pacific	No	Yes
Lbul_KH1110_08_40	N. Pacific	KH1110_08	RMNH.MOL.346523	SAMN20293172	MZ542629	6542126	2871694	43.9	96.7	206.91	Indo-Pacific	No	Yes
Lbul_KH1110_15_01	S. Pacific	KH1110_15	RMNH.MOL.346524	SAMN20293173	MZ542630	10670389	2387520	22.4	97.9	174.54	Pacific	No	Yes
Lbul_KH1110_15_03	S. Pacific	KH1110_15	RMNH.MOL.346525	SAMN20293174	MZ542631	12318189	5340424	43.4	95.8	379.31	Pacific	Yes	Yes
Lbul_KH1110_15_05	S. Pacific	KH1110_15	RMNH.MOL.346526	SAMN20293175	MZ542632	8232996	3610486	43.9	95.1	254.63	Pacific	No	Yes
Lbul_KH1110_15_06	S. Pacific	KH1110_15	RMNH.MOL.346527	SAMN20293176	MZ542633	9424493	4305134	45.7	96.1	307.82	Pacific	No	Yes
Lbul_KH1110_15_07	S. Pacific	KH1110_15	RMNH.MOL.346528	SAMN20293177	MZ542634	11400718	4953148	43.4	96.1	354.77	Pacific	No	Yes
Lbul_KH1110_15_08	S. Pacific	KH1110_15	RMNH.MOL.346529	SAMN20293178	MZ542635	11182753	2396388	21.4	98	175.15	Pacific	No	Yes
Lbul_KH1110_15_09	S. Pacific	KH1110_15	RMNH.MOL.346530	SAMN20293179	MZ542636	8452428	3633690	43.0	96.9	262.21	Pacific	No	Yes
Lbul_KH1110_15_12	S. Pacific	KH1110_15	RMNH.MOL.346531	SAMN20293180	MZ542637	8923885	3788590	42.5	96	271	Pacific	No	Yes
Lbul_KH1110_15_13	S. Pacific	KH1110_15	RMNH.MOL.346532	SAMN20293181	MZ542638	7608276	3231348	42.5	95.6	230.57	Pacific	No	Yes

Table S1 Continued.

Specimen ID	Location	Station	Museum voucher code	SRA	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth	Lineage	SNAPP/ stairway plot	Morpho-metrics
Lbul_KH1110_15_14	S. Pacific	KH1110_15	RMNH.MOL.346533	SAMN20293182	MZ542639	13138144	5346350	40.7	95.8	380.33	Pacific	Yes	Yes
Lbul_KH1110_15_15	S. Pacific	KH1110_15	RMNH.MOL.346534	SAMN20293183	MZ542640	8369910	3208036	38.3	95.7	228.41	Pacific	No	Yes
Lbul_KH1110_18_03	S. Pacific	KH1110_18	RMNH.MOL.346535	SAMN20293184	MZ542641	14673417	5511052	37.6	94.8	387.97	Pacific	No	Yes
Lbul_KH1110_18_04	S. Pacific	KH1110_18	RMNH.MOL.346536	SAMN20293185	MZ542642	10106228	4490020	44.4	95.8	320.35	Pacific	No	Yes
Lbul_KH1110_18_07	S. Pacific	KH1110_18	RMNH.MOL.346537	SAMN20293186	MZ542643	5113657	1781382	34.8	95.5	125.97	Pacific	No	Yes
Lbul_KH1110_18_08	S. Pacific	KH1110_18	RMNH.MOL.346538	SAMN20293187	MZ542644	11090186	4495290	40.5	95.1	318.65	Pacific	No	Yes
Lbul_KH1110_18_09	S. Pacific	KH1110_18	RMNH.MOL.346539	SAMN20293188	MZ542645	8466526	3789158	44.8	95.9	269.91	Pacific	No	Yes
Lbul_KH1110_18_10	S. Pacific	KH1110_18	RMNH.MOL.346540	SAMN20293189	MZ542646	14049538	3090806	22.0	97.3	223.77	Pacific	No	Yes
Lbul_KH1110_18_11	S. Pacific	KH1110_18	RMNH.MOL.346541	SAMN20293190	MZ542647	7233698	2840436	39.3	96.9	204.39	Pacific	No	Yes
Lbul_KH1110_18_12	S. Pacific	KH1110_18	RMNH.MOL.346542	SAMN20293191	MZ542648	9515292	3936516	41.4	96.7	283.43	Pacific	No	Yes
Lbul_KH1110_18_13	S. Pacific	KH1110_18	RMNH.MOL.346543	SAMN20293192	MZ542649	8461723	3314520	39.2	95.5	235.88	Pacific	No	Yes
Lbul_KH1110_18_14	S. Pacific	KH1110_18	RMNH.MOL.346544	SAMN20293193	MZ542650	8886176	1962366	22.1	97.9	142.89	Pacific	No	Yes
Lbul_KH1110_18_15	S. Pacific	KH1110_18	RMNH.MOL.346545	SAMN20293194	MZ542651	14215968	5793088	40.8	95.1	407.78	Pacific	Yes	Yes
Lbul_KH1110_21_01	S. Pacific	KH1110_21	RMNH.MOL.346546	SAMN20293195	MZ542652	9021724	3506922	38.9	95.5	248.94	Pacific	No	Yes
Lbul_KH1110_21_04	S. Pacific	KH1110_21	RMNH.MOL.346547	SAMN20293196	MZ542653	9389875	4332232	46.1	95.2	306.91	Pacific	No	Yes
Lbul_KH1110_21_05	S. Pacific	KH1110_21	RMNH.MOL.346548	SAMN20293197	MZ542654	8475021	3624448	42.8	96.4	261.1	Pacific	No	Yes
Lbul_KH1110_21_07	S. Pacific	KH1110_21	RMNH.MOL.346549	SAMN20293198	MZ542655	6883168	3213776	46.7	96.5	231.56	Pacific	No	Yes
Lbul_KH1110_21_08	S. Pacific	KH1110_21	RMNH.MOL.346550	SAMN20293199	MZ542656	15141791	5981748	39.5	96.6	429.6	Pacific	No	Yes
Lbul_KH1110_21_09	S. Pacific	KH1110_21	RMNH.MOL.346551	SAMN20293200	MZ542657	8071325	3302934	40.9	96.4	237.17	Pacific	No	Yes
Lbul_KH1110_21_10	S. Pacific	KH1110_21	RMNH.MOL.346552	SAMN20293201	MZ542658	9566905	3929578	41.1	96.3	282.34	Pacific	No	Yes
Lbul_KH1110_21_11	S. Pacific	KH1110_21	RMNH.MOL.346553	SAMN20293202	MZ542659	17270860	3471166	20.1	98.2	253.47	Pacific	No	Yes
Lbul_KH1110_21_13	S. Pacific	KH1110_21	RMNH.MOL.346554	SAMN20293203	MZ542660	9961733	3973154	39.9	95.1	281.11	Pacific	No	Yes
Lbul_KH1110_21_14	S. Pacific	KH1110_21	RMNH.MOL.346555	SAMN20293204	MZ542661	15279774	5074560	33.2	93.6	351.79	Pacific	Yes	Yes
Lbul_KH1110_21_15	S. Pacific	KH1110_21	RMNH.MOL.346556	SAMN20293205	MZ542662	9717633	3951344	40.7	95.6	281.55	Pacific	No	Yes
Lbul_KOK1703_03_05	N. Pacific	KOK1703_03	RMNH.MOL.340360	SAMN20293206	MZ542663	12181573	4085842	33.5	81.2	232.91	Pacific	Yes	Yes
Lbul_KOK1703_03_06	N. Pacific	KOK1703_03	RMNH.MOL.340361	SAMN20293207	MZ542664	13448333	6371976	47.4	91	416.26	Pacific	Yes	Yes
Lbul_KOK1703_03_07	N. Pacific	KOK1703_03	RMNH.MOL.340362	SAMN20293208	MZ542665	8290806	3949554	47.6	92.3	262.89	Pacific	Yes	Yes
Lbul_KOK1703_03_08	N. Pacific	KOK1703_03	RMNH.MOL.340363	SAMN20293209	MZ542666	9056301	4403198	48.6	92	292.76	Indo-Pacific	Yes	Yes
Lbul_KOK1703_03_09	N. Pacific	KOK1703_03	RMNH.MOL.340364	SAMN20293210	MZ542667	3780542	1159706	30.7	80.4	64.59	Indo-Pacific	No	Yes
Lbul_KOK1703_03_10	N. Pacific	KOK1703_03	RMNH.MOL.340365	SAMN20293211	MZ542668	11431899	5652996	49.4	86.4	349.51	Indo-Pacific	Yes	Yes
Lbul_KOK1703_03_11	N. Pacific	KOK1703_03	RMNH.MOL.340366	SAMN20293212	MZ542669	5045107	2553772	50.6	92.6	170.77	Pacific	Yes	Yes
Lbul_KOK1703_03_13	N. Pacific	KOK1703_03	RMNH.MOL.340367	SAMN20293213	MZ542670	6526049	3427222	52.5	91.9	227.7	Indo-Pacific	Yes	Yes
Lbul_KOK1703_03_14	N. Pacific	KOK1703_03	RMNH.MOL.340368	SAMN20293214	MZ542671	9337697	2959712	31.7	87.2	183.21	Pacific	Yes	Yes
Lbul_KOK1703_03_15	N. Pacific	KOK1703_03	RMNH.MOL.340369	SAMN20293215	MZ542672	9813374	5101536	52.0	91.4	336.6	Indo-Pacific	Yes	Yes
Lbul_NIC2_S1C3_01	Atlantic	NIC2_S1C3	RMNH.MOL.346557	SAMN20293216	MZ542673	17269646	5816130	33.7	96.5	418.65	Atlantic	No	Yes
Lbul_NIC2_S1C3_02	Atlantic	NIC2_S1C3	RMNH.MOL.346558	SAMN20293217	MZ542674	11362364	4261032	37.5	96.2	305.98	Atlantic	No	Yes
Lbul_NIC2_S1C3_03	Atlantic	NIC2_S1C3	RMNH.MOL.346559	SAMN20293218	MZ542675	13073572	2907762	22.2	98.2	214.8	Atlantic	No	Yes

TABLE S1 Continued.

Specimen ID	Location	Station	Museum voucher code	SRA	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth	Lineage	SNAPP/ stairway plot	Morpho-metrics
Lbul_NIC2_S1C3_07	Atlantic	NIC2_S1C3	RMNH.MOL.346560	SAMN20293219	MZ542676	19015434	3526366	18.5	98.3	258.02	Atlantic	No	Yes
Lbul_NIC2_S1C3_08	Atlantic	NIC2_S1C3	RMNH.MOL.346561	SAMN20293220	MZ542677	12518723	4716858	37.7	95	333.98	Atlantic	No	Yes
Lbul_NIC2_S1C3_10	Atlantic	NIC2_S1C3	RMNH.MOL.346562	SAMN20293221	MZ542678	11557090	4428220	38.3	96.1	317.52	Atlantic	No	Yes
Lbul_NIC2_S1C3_11	Atlantic	NIC2_S1C3	RMNH.MOL.346563	SAMN20293222	MZ542679	9691816	3957806	40.8	97.1	288.39	Atlantic	No	Yes
Lbul_NIC2_S1C3_12	Atlantic	NIC2_S1C3	RMNH.MOL.346564	SAMN20293223	MZ542680	12914232	5011224	38.8	96.2	359.34	Atlantic	No	Yes
Lbul_NIC2_S1C3_13	Atlantic	NIC2_S1C3	RMNH.MOL.346565	SAMN20293224	MZ542681	12934158	2699656	20.9	97.8	197.4	Atlantic	No	Yes
Lbul_NIC2_S1C3_14	Atlantic	NIC2_S1C3	RMNH.MOL.346566	SAMN20293225	MZ542682	6773313	2628106	38.8	96.8	190	Atlantic	No	Yes
Lbul_NIC2_S1C3_15	Atlantic	NIC2_S1C3	RMNH.MOL.346567	SAMN20293226	MZ542683	15422685	4556206	29.5	94.2	317.07	Atlantic	Yes	Yes
Lbul_NIC2_S9C3_01	Atlantic	NIC2_S9C3	RMNH.MOL.346568	SAMN20293227	MZ542684	9518473	3873588	40.7	96.9	280.88	Atlantic	No	Yes
Lbul_NIC2_S9C3_02	Atlantic	NIC2_S9C3	RMNH.MOL.346569	SAMN20293228	MZ542685	12754807	2595230	20.3	98.4	190.72	Atlantic	No	Yes
Lbul_NIC2_S9C3_03	Atlantic	NIC2_S9C3	RMNH.MOL.346570	SAMN20293229	MZ542686	12515237	2553328	20.4	98.2	186.96	Atlantic	No	Yes
Lbul_NIC2_S9C3_04	Atlantic	NIC2_S9C3	RMNH.MOL.346571	SAMN20293230	MZ542687	11723092	2511120	21.4	97.9	183.6	Atlantic	No	Yes
Lbul_NIC2_S9C3_05	Atlantic	NIC2_S9C3	RMNH.MOL.346572	SAMN20293231	MZ542688	6621308	2916484	44.0	96.1	209.55	Atlantic	No	Yes
Lbul_NIC2_S9C3_06	Atlantic	NIC2_S9C3	RMNH.MOL.346573	SAMN20293232	MZ542689	16042144	3317304	20.7	98.2	243.47	Atlantic	No	Yes
Lbul_NIC2_S9C3_07	Atlantic	NIC2_S9C3	RMNH.MOL.346574	SAMN20293233	MZ542690	11817232	5046408	42.7	95.4	358.51	Atlantic	No	Yes
Lbul_NIC2_S9C3_08	Atlantic	NIC2_S9C3	RMNH.MOL.346575	SAMN20293234	MZ542691	8288640	3383882	40.8	96.8	245.5	Atlantic	No	Yes
Lbul_NIC2_S9C3_09	Atlantic	NIC2_S9C3	RMNH.MOL.346576	SAMN20293235	MZ542692	9531410	4113674	43.2	96.4	296.07	Atlantic	No	Yes
Lbul_NIC2_S9C3_10	Atlantic	NIC2_S9C3	RMNH.MOL.346577	SAMN20293236	MZ542693	8715280	3681980	42.2	97	266.83	Atlantic	No	Yes
Lbul_NIC8_S5C3_01	Atlantic	NIC8_S5C3	RMNH.MOL.346578	SAMN20293237	MZ542694	7064864	2931840	41.5	96.4	211.49	Atlantic	No	Yes
Lbul_NIC8_S5C3_03	Atlantic	NIC8_S5C3	RMNH.MOL.346579	SAMN20293238	MZ542695	8865106	3757808	42.4	96.4	271.51	Atlantic	No	Yes
Lbul_NIC8_S5C3_04	Atlantic	NIC8_S5C3	RMNH.MOL.346580	SAMN20293239	MZ542696	14846520	5444886	36.7	94.6	383.93	Atlantic	Yes	Yes
Lbul_NIC8_S5C3_06	Atlantic	NIC8_S5C3	RMNH.MOL.346581	SAMN20293240	MZ542697	12715480	5027460	39.5	96.4	362.09	Atlantic	No	Yes
Lbul_NIC8_S5C3_07	Atlantic	NIC8_S5C3	RMNH.MOL.346582	SAMN20293241	MZ542698	11893813	4701916	39.5	96.6	339.21	Atlantic	No	Yes
Lbul_NIC8_S5C3_08	Atlantic	NIC8_S5C3	RMNH.MOL.346583	SAMN20293242	MZ542699	16832853	6087524	36.2	95.6	434.87	Atlantic	No	Yes
Lbul_NIC8_S5C3_09	Atlantic	NIC8_S5C3	RMNH.MOL.346584	SAMN20293243	MZ542700	10162231	2142102	21.1	97.2	155.04	Atlantic	No	Yes
Lbul_NIC8_S5C3_11	Atlantic	NIC8_S5C3	RMNH.MOL.346585	SAMN20293244	MZ542701	8608429	2917080	33.9	95.4	206.51	Atlantic	No	Yes
Lbul_NIC8_S5C3_12	Atlantic	NIC8_S5C3	RMNH.MOL.346586	SAMN20293245	MZ542702	14259468	5927612	41.6	96.2	424.53	Atlantic	Yes	Yes
Lbul_NIC8_S5C3_13	Atlantic	NIC8_S5C3	RMNH.MOL.346587	SAMN20293246	MZ542703	6800160	2862764	42.1	97.1	208.5	Atlantic	No	Yes
Lbul_NIC8_S5C3_14	Atlantic	NIC8_S5C3	RMNH.MOL.346588	SAMN20293247	MZ542704	8123312	3435518	42.3	96.5	248.52	Atlantic	No	Yes
Lbul_SN105_08_01	Indian	SN105_08	RMNH.MOL.346589	SAMN20293248	MZ542705	7093142	3240374	45.7	95.4	229.28	Indo-Pacific	Yes	Yes
Lbul_SN105_08_02	Indian	SN105_08	RMNH.MOL.346590	SAMN20293249	MZ542706	11961694	5084258	42.5	94.9	357.19	Indo-Pacific	Yes	Yes
Lbul_SN105_08_03	Indian	SN105_08	RMNH.MOL.346591	SAMN20293250	MZ542707	6389774	2970202	46.5	96.1	212.39	Indo-Pacific	No	Yes
Lbul_SN105_08_04	Indian	SN105_08	RMNH.MOL.346592	SAMN20293251	MZ542708	8352779	3700260	44.3	96	263.27	Indo-Pacific	No	Yes
Lbul_SN105_08_06	Indian	SN105_08	RMNH.MOL.346593	SAMN20293252	MZ542709	9092007	3558366	39.1	95.2	251.92	Indo-Pacific	No	Yes
Lbul_SN105_08_07	Indian	SN105_08	RMNH.MOL.346594	SAMN20293253	MZ542710	7707444	3192652	41.4	95.9	227.93	Indo-Pacific	No	No
Lbul_SN105_08_08	Indian	SN105_08	RMNH.MOL.346595	SAMN20293254	MZ542711	7161433	2876190	40.2	95.8	205.02	Indo-Pacific	No	Yes
Lbul_SN105_08_09	Indian	SN105_08	RMNH.MOL.346596	SAMN20293255	MZ542712	4550748	1496494	32.9	94.6	104.42	Indo-Pacific	No	Yes

TABLE S1 Continued.

Specimen ID	Location	Station	Museum voucher code	SRA	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth	Lineage	SNAPP/ stairway plot	Morpho-metrics
Lbul_SN105_08_10	Indian	SN105_08	RMNH.MOL.346597	SAMN20293256	MZ542713	12826076	5102068	39.8	93.5	352.35	Indo-Pacific	Yes	Yes
Lbul_SN105_08_11	Indian	SN105_08	RMNH.MOL.346598	SAMN20293257	MZ542714	9254717	3508730	37.9	95.8	249.68	Indo-Pacific	No	Yes
Lbul_SN105_08_12	Indian	SN105_08	RMNH.MOL.346599	SAMN20293258	MZ542715	6894729	2726754	39.5	95.4	193.15	Indo-Pacific	No	Yes
Lbul_SO255_143_02	S. Pacific	SO255_143	RMNH.MOL.346600	SAMN20293259	MZ542716	12748730	4850524	38.0	95.4	343.78	Pacific	No	Yes
Lbul_SO255_143_05	S. Pacific	SO255_143	RMNH.MOL.346601	SAMN20293260	MZ542717	7631666	3219798	42.2	95.9	230.38	Pacific	No	Yes
Lbul_SO255_143_06	S. Pacific	SO255_143	RMNH.MOL.346602	SAMN20293261	MZ542718	8555574	3767896	44.0	95.4	267.58	Pacific	No	Yes
Lbul_SO255_143_07	S. Pacific	SO255_143	RMNH.MOL.346603	SAMN20293262	MZ542719	9484173	3953554	41.7	95.1	277.99	Pacific	No	Yes
Lbul_SO255_143_08	S. Pacific	SO255_143	RMNH.MOL.346604	SAMN20293263	MZ542720	8463015	3580560	42.3	96.5	257.86	Pacific	No	Yes
Lbul_SO255_143_09	S. Pacific	SO255_143	RMNH.MOL.346605	SAMN20293264	MZ542721	5931457	2563858	43.2	96.6	184.88	Pacific	No	Yes
Lbul_SO255_143_10	S. Pacific	SO255_143	RMNH.MOL.346606	SAMN20293265	MZ542722	6958261	2782606	40.0	95.4	197.82	Pacific	No	Yes
Lbul_SO255_143_12	S. Pacific	SO255_143	RMNH.MOL.346607	SAMN20293266	MZ542723	6482019	2605690	40.2	96.6	187.78	Pacific	No	Yes
Lbul_SO255_143_13	S. Pacific	SO255_143	RMNH.MOL.346608	SAMN20293267	MZ542724	24974402	6477318	25.9	94	449.4	Pacific	Yes	Yes
Lbul_SO255_143_14	S. Pacific	SO255_143	RMNH.MOL.346609	SAMN20293268	MZ542725	8105002	3267556	40.3	96.2	234.35	Pacific	No	Yes
Lbul_SO255_143_15	S. Pacific	SO255_143	RMNH.MOL.346610	SAMN20293269	MZ542726	9150047	3730884	40.8	96.2	267.61	Pacific	No	Yes

1.2. Genetic lineage assignment for North Pacific sites

Specimen	Station	Population
Lbul_KH1110_02_09	KH1110_02	Indo-Pacific
Lbul_KH1110_02_10	KH1110_02	Indo-Pacific
Lbul_KH1110_02_13	KH1110_02	Indo-Pacific
Lbul_KH1110_05_02	KH1110_05	Indo-Pacific
Lbul_KH1110_05_04	KH1110_05	Indo-Pacific
Lbul_KH1110_05_05	KH1110_05	Indo-Pacific
Lbul_KH1110_05_06	KH1110_05	Indo-Pacific
Lbul_KH1110_05_07	KH1110_05	Indo-Pacific
Lbul_KH1110_05_08	KH1110_05	Indo-Pacific
Lbul_KH1110_05_10	KH1110_05	Indo-Pacific
Lbul_KH1110_05_12	KH1110_05	Indo-Pacific
Lbul_KOK1703_03_08	KOK1703_03	Indo-Pacific
Lbul_KOK1703_03_09	KOK1703_03	Indo-Pacific
Lbul_KOK1703_03_10	KOK1703_03	Indo-Pacific
Lbul_KOK1703_03_13	KOK1703_03	Indo-Pacific
Lbul_KOK1703_03_15	KOK1703_03	Indo-Pacific
Lbul_KH1110_02_03	KH1110_02	Pacific
Lbul_KH1110_02_05	KH1110_02	Pacific
Lbul_KH1110_02_06	KH1110_02	Pacific
Lbul_KH1110_02_07	KH1110_02	Pacific
Lbul_KH1110_02_08	KH1110_02	Pacific
Lbul_KH1110_02_11	KH1110_02	Pacific
Lbul_KH1110_02_12	KH1110_02	Pacific
Lbul_KH1110_02_14	KH1110_02	Pacific
Lbul_KH1110_05_01	KH1110_05	Pacific
Lbul_KH1110_05_09	KH1110_05	Pacific
Lbul_KH1110_05_11	KH1110_05	Pacific
Lbul_KOK1703_03_05	KOK1703_03	Pacific
Lbul_KOK1703_03_06	KOK1703_03	Pacific
Lbul_KOK1703_03_07	KOK1703_03	Pacific
Lbul_KOK1703_03_11	KOK1703_03	Pacific
Lbul_KOK1703_03_14	KOK1703_03	Pacific

2. Methods protocols

2.1. Modifications to DNA extraction protocol

Modified DNA extraction protocol with E.Z.N.A Insect Kit (Omega Bio-tek) for extracting DNA from shelled pteropods

1. Soak the complete individual in nuclease-free water for 30 minutes.

2. Transfer individual into 2 ml screw cap grinding tube containing:

- Precellys 1.4mm zirconium oxide bashing beads
- 400 µl CTL buffer
- 28.5 µl Proteinase K solution

3. Grind individual for 20 seconds at a frequency of 30 per second.

4. Incubate sample at 60 °C for 20 minutes.

5. Repeat step *3* and incubate sample at 60°C for 20 minutes.

6. Add 400 µl chloroform:isoamyl alcohol, vortex for 10 seconds.

7. Centrifuge at 13,500 g for 18 minutes at room temperature.

8. Prepare the HiBind DNA Mini Column:

- Assemble to collection tube
- Add 100 µl 3 M NaOH (equilibration buffer) on filter
- Incubate for 4 minutes and centrifuge at 17,000 g for 20 seconds.

9. Transfer upper aqueous phase (~320 µl) from step 7 to a new 1.5 ml eppendorf tube. Avoid the white milky surface.
10. Add one volume CBL buffer.
11. Add one volume 100% EtOH and vortex for 10 seconds.
12. Transfer 750 µl of the mixture to the equilibrated HiBind DNA Mini Column and centrifuge at 10,000 g for 1 minute. Discard the flow-through.
13. Repeat step 12 until all of the mixture has been used and insert DNA Mini Column in a new collection tube
14. Warm-up the Elution Buffer at 70 °C.
15. Add 500 µl HBC buffer and centrifuge at 10,000 g for 30 seconds. Discard the flow-through.
16. Add 500 µl DNA Wash Buffer and centrifuge at 10,000 g for 1 minute. Discard the flow-through.
17. Repeat step 16.
18. Re-insert HiBind DNA Mini Column into new collection tube, centrifuge at 15,000 g for 2 minutes.
19. Re-insert HiBind DNA minicolumn into new eppendorf tube.
20. Add 40 µl of pre-heated Elution Buffer directly to the center of the Mini Column membrane without touching it.
21. Wait for 2 minutes and then centrifuge at 10,000 g for 1 minute.
22. Repeat steps 20 and 21 for a final elution volume of 80 µl.

2.2. Target capture protocol

All DNA libraries were fragmented to an average size of 300 bp by sonication and were prepared using the NEXTflex™ Rapid Pre-Capture Combo Kit (Bioo Scientific, Austin, TX, USA), including a step of single adapter indexing of each library. Libraries were cleaned-up and amplified separately for 8 cycles. 26 to 27 libraries were pooled in equimolar ratios for each capture reaction. We increased the efficiency of the hybridisation and aimed to maximise the number of on-target captured sequences by performing the capture reaction twice, splitting the total amount of baits required for one reaction in two. The first round of hybridisation was performed using 4 µl of baits for each reaction. The reaction was performed over three days at a temperature of 60 °C in order to maximise the specificity. Capture was performed consecutively using DYNAbeads MyOne Streptavidin C1 beads (Invitrogen) to bind the hybridised targets during 30 min at 65°C. The captured DNA was amplified by PCR for 8 cycles using KAPA HiFi HotStart ReadyMix (Kapa Biosystems). A second round of hybridisation was conducted using 1.5 µl of baits for each of the two pools, followed by a second capture and 6 or 7 more cycles of post-capture PCR. Finally, three pools each were mixed together in equimolar concentrations and sequenced on one NextSeq 550 (Illumina) 2x150 bp high-output kit v.2.

3. Data processing steps

3.1. Flowchart of filtering protocols used for analyses

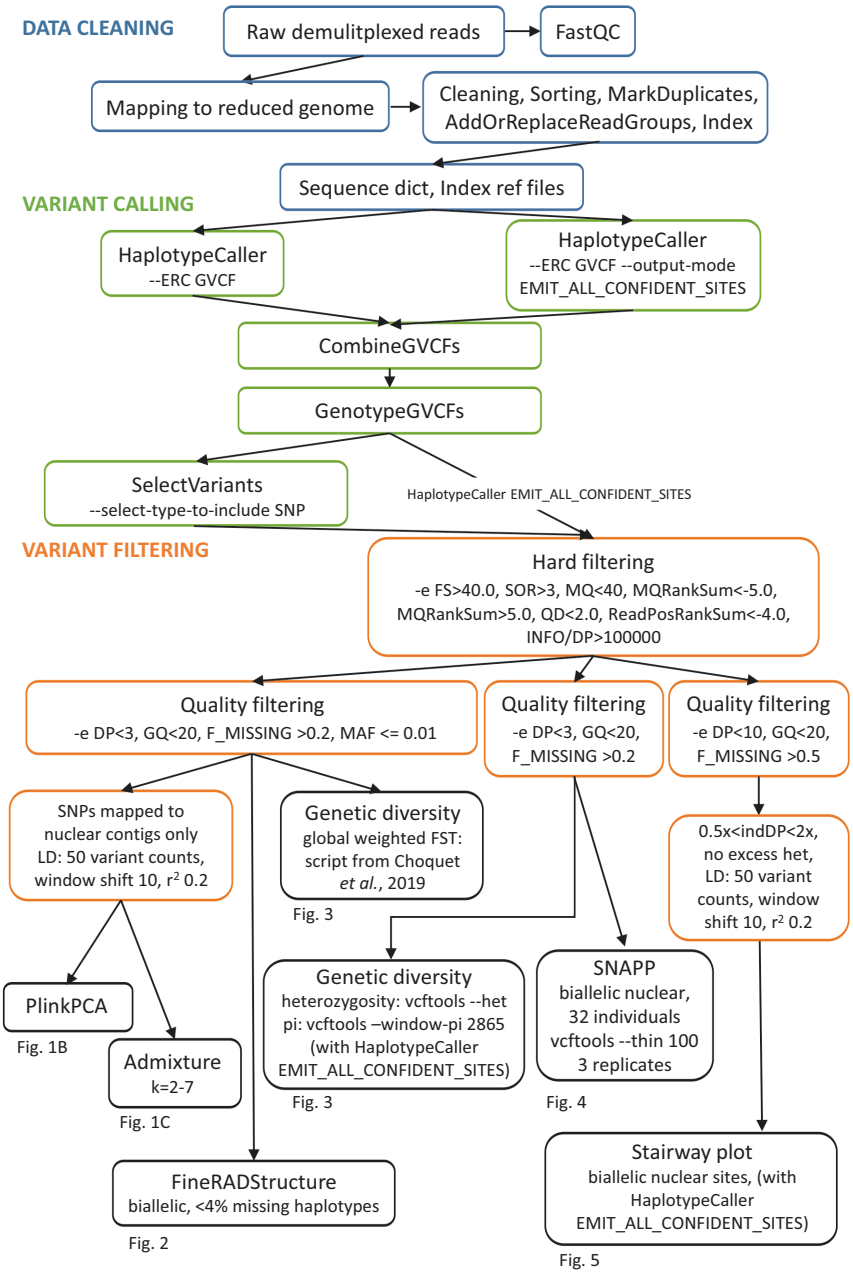


Figure S1 Single nucleotide polymorphism (SNP) calling pipeline and settings for downstream genetic analyses.

3.2. Landmark positions on shell image

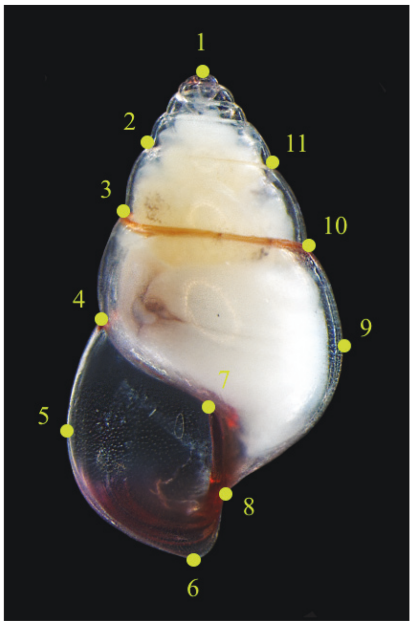


Figure S2 Eleven landmarks used to quantify shell shape variation in *Limacina bulimoides*. 1: On the apex. 2: On the left border of the profile of the shell at the end of the upper suture of the penultimate whorl. 3: On the left border of the profile of the shell at the end of the upper suture of the last whorl. 4: On the left border of the profile of the shell at the end of the suture of the last whorl. 5: On the left outermost point of the apertural lip. 6: On the lowest point of the whole shell. 7: Where the parietal wall meets the columellar whorl. 8: Where the right outline of the last whorl meets the inner aperture. 9: On the outermost point on the right outline of the last whorl. 10: On the right outline of the shell at the end of the upper suture of the last whorl. 11: On the right border of the outline of the shell at the end of the upper suture of the penultimate whorl.

4. Additional results plots/tables

4.1. Genetic analyses

4.1.1. Recovered single nucleotide polymorphisms and contigs from each lineage

Table S2 Number of single nucleotide polymorphism (SNPs) and contigs recovered after hard filtering implemented in GATK4.1.7.0 VariantFiltration using the following settings: QualByDepth <2.0, FisherStrand >60.0, RMSMappingQuality <5.0, MQRankSumTest <-5.0 and ReadPositionRankSum <-5.0. The SNPs were identified to their corresponding category of targeted regions, namely candidate biomineralisation genes (Biomin.), conserved pteropod orthologues (Ortholog.), mitochondrial (Mt genes), nuclear 28S, and other coding and non-coding regions. Numbers in brackets represent the total number of targets in that category on the set of target probes designed for *Limacina bulimoides*. SNPs in common were calculated for all three lineages, or for the intersection (∩) between the different pairs of lineages.

Category	Biomin. (35)	Ortholog. (643)	Mt genes (10)	28S (1)	Coding (2169)	Non-coding (42)	Total (2900)
SNPs							
Global	18,716	328,298	262	23	1,246,720	35,363	1,629,382
Atlantic	13,168	212,689	249	18	779,546	25,619	1,031,289
Indo-Pacific	8,596	129,211	142	7	461,335	14,267	613,558
Pacific	13,509	225,332	176	10	838,466	20,693	1,098,186
Contigs							
Global	35	633	2	1	2,089	42	2,802
Atlantic	35	633	2	1	2,089	42	2,802
Indo-Pacific	33	629	2	1	2,072	32	2,769
Pacific	33	630	2	1	2,071	36	2,773
Unique SNPs							
Atlantic	10,134	165,263	228	11	599,126	19,104	793,866
Indo-Pacific	5,851	86,468	51	3	299,521	8,508	400,402
Pacific	10,389	175,068	81	5	645,308	14,248	845,099
Common SNPs							
3 lineages	633	10,541	11	2	42,952	2,057	56,196
Atlantic∩Indo-Pacific	1,013	14,682	3	0	53,062	1,886	70,646
Atlantic∩Pacific	1,388	22,203	7	1	84,406	2,572	110,577
Indo-Pacific∩Pacific	1,099	17,520	77	3	65,800	1,816	86,315

4.1.2. COI haplotype network

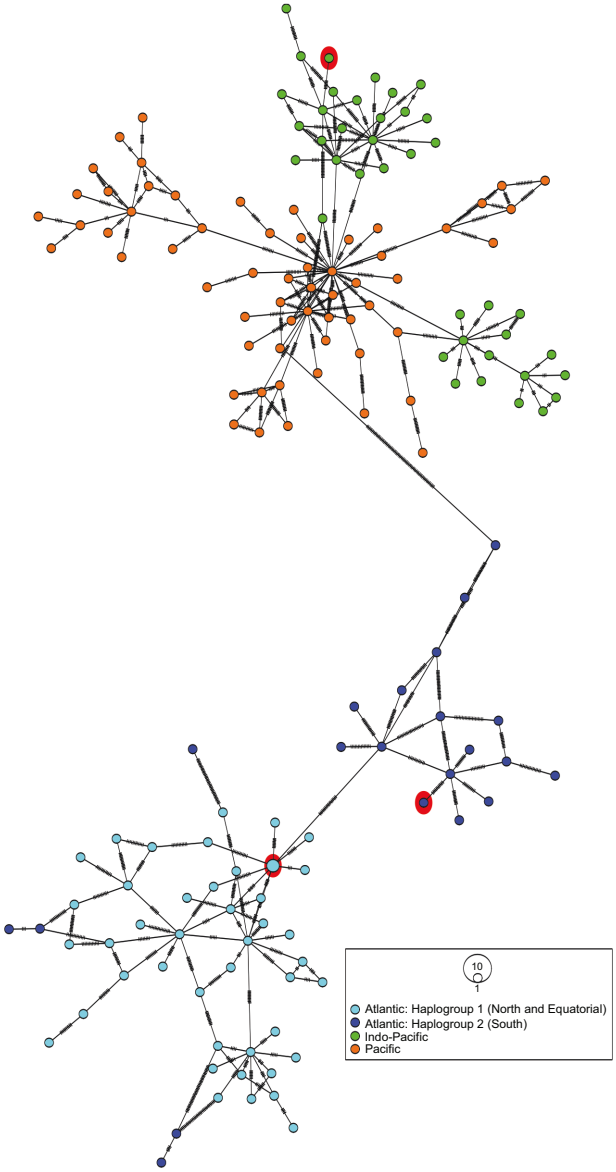


FIGURE S3 Minimum spanning network for the mitochondrial cytochrome c oxidase I (COI) gene fragment obtained from 161 individuals of *Limacina bulimoides*. The size of the filled circles represents the number of individuals with each haplotype, with the smallest circles representing one individual with that haplotype, while colour represents previously documented COI haplogroups in the Atlantic (Haplogroup 1 and 2 in the North/Equatorial and South Atlantic, respectively) as well as the Indo-Pacific and Pacific lineages. The three sequences from (Choo et al., 2021) are indicated with a red highlight around their haplotype (Genbank accessions: MN952611, MN952937, MN952944).

4.2. Morphometric analyses

4.2.1. Variation explained by relative warps

Table S3 Singular values and percentage of shell shape variation explained by each of the six repeatable relative warps.

RW	Singular values	Variation explained (%)
1	0.44584	51.05
2	0.26593	18.16
3	0.17783	8.12
5	0.13241	4.50
10	0.06717	1.16
11	0.05433	0.76
Total		83.75

4.2.2. Canonical variate analysis by lineage

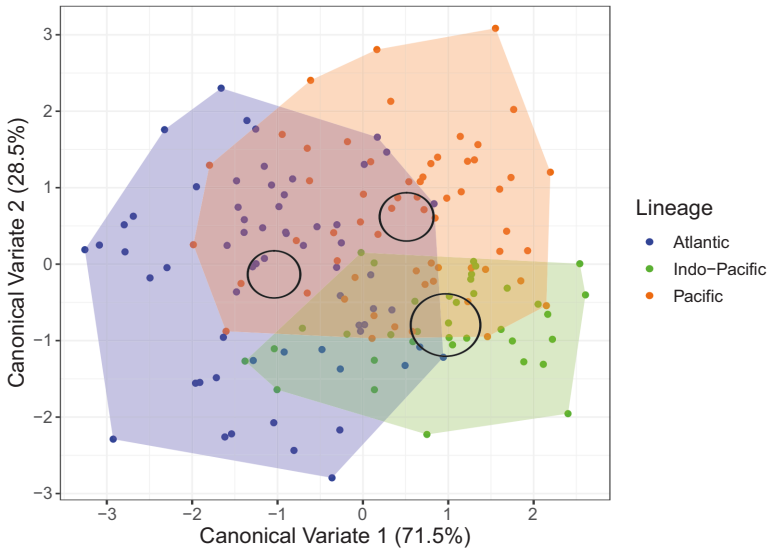


Figure S4 Canonical variate analysis for differences in mean shell shape based on the six repeatable relative warps calculated for 159 *Limacina bulimoides* individuals that were assigned to three lineages: Atlantic, Indo-Pacific and Pacific. The 95% confidence regions for the mean of each lineage and the canonical variate space occupied by each lineage are represented by the circled regions and coloured polygons, respectively.

4.2.3. Relative warp analysis by location

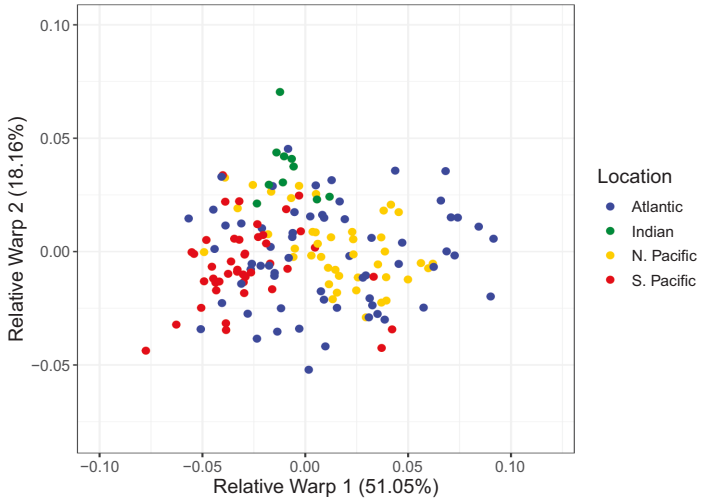


FIGURE S5 Shell shape variation of 159 *Limacina bulimoides*, categorised into four ocean basins (Atlantic, Indian, North Pacific and South Pacific). Shell shape variation is visualised with relative warps 1 and 2, which explain 51.05% and 18.16% of shell shape variation, respectively.

4.2.4. Canonical variate analysis by location

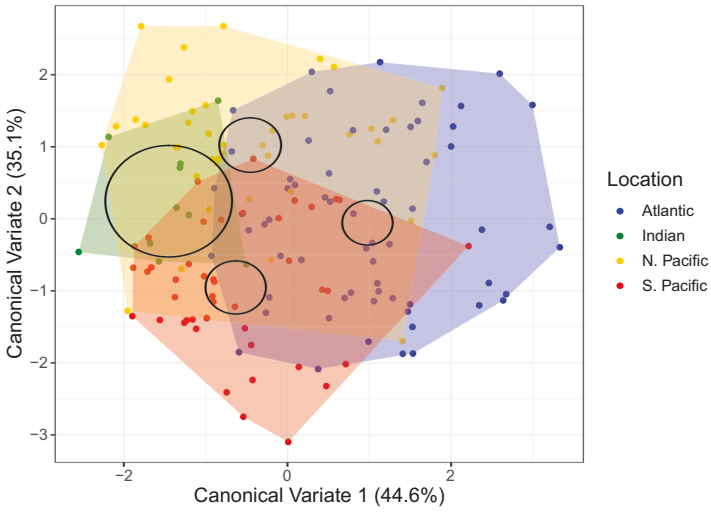
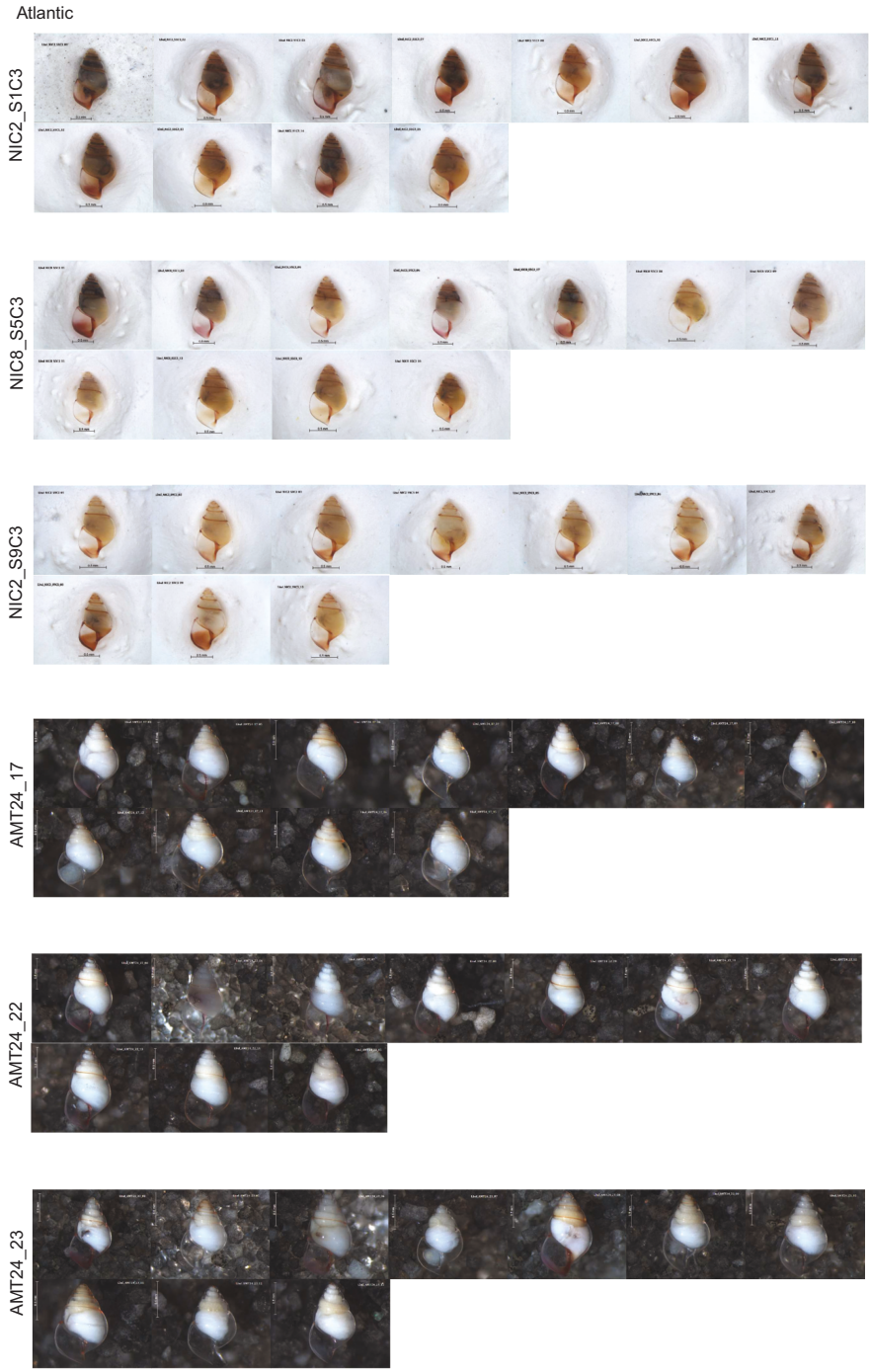


FIGURE S6 Canonical variate analysis for differences in mean shell shape based on the six repeatable relative warps calculated for 159 *Limacina bulimoides* individuals that were assigned into four locations: Atlantic, Indo-Pacific and Pacific. The 95% confidence regions for the mean of each location and the canonical variate space occupied by each location are represented by the circled regions and coloured polygons, respectively.

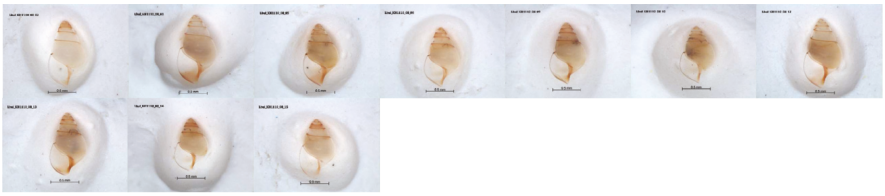
4.3. Shell images

4.3.1. Overview of all shell images

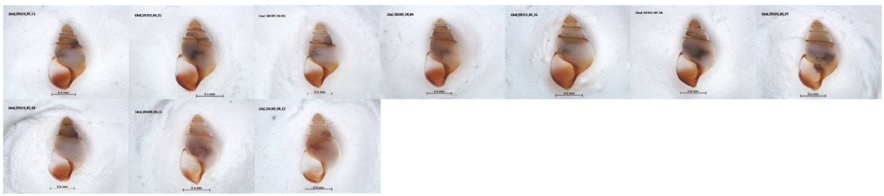


Indo-Pacific

KH1110_08

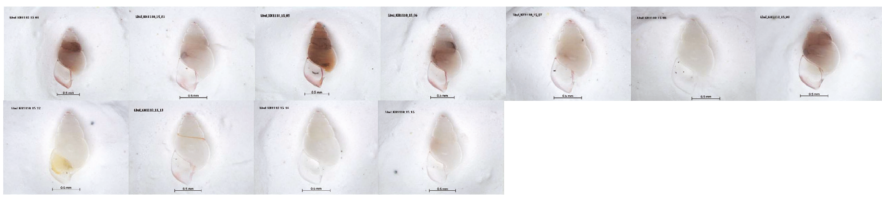


SN105_08

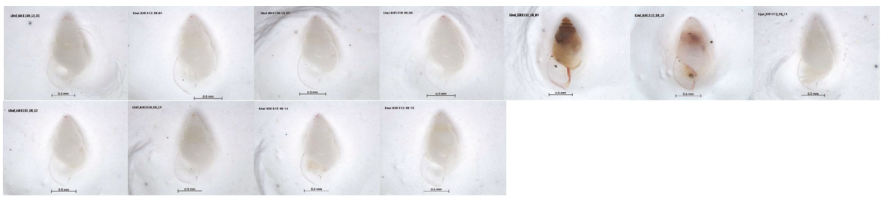


Pacific

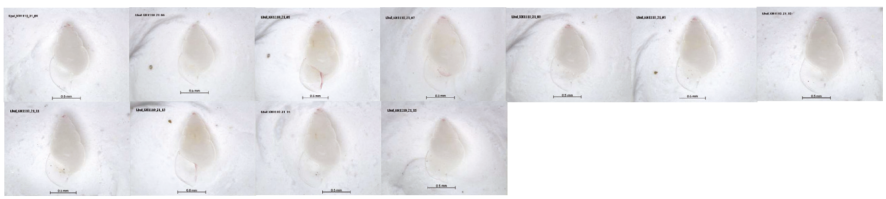
KH1110_15



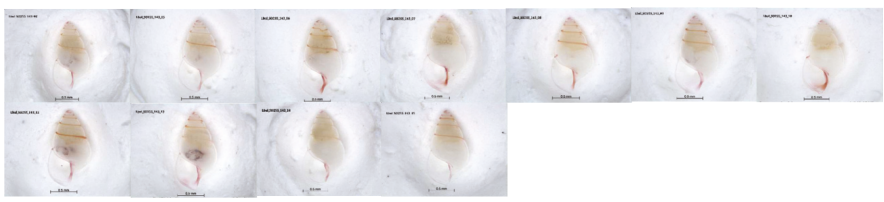
KH1110_18



KH1110_21



SO255_143



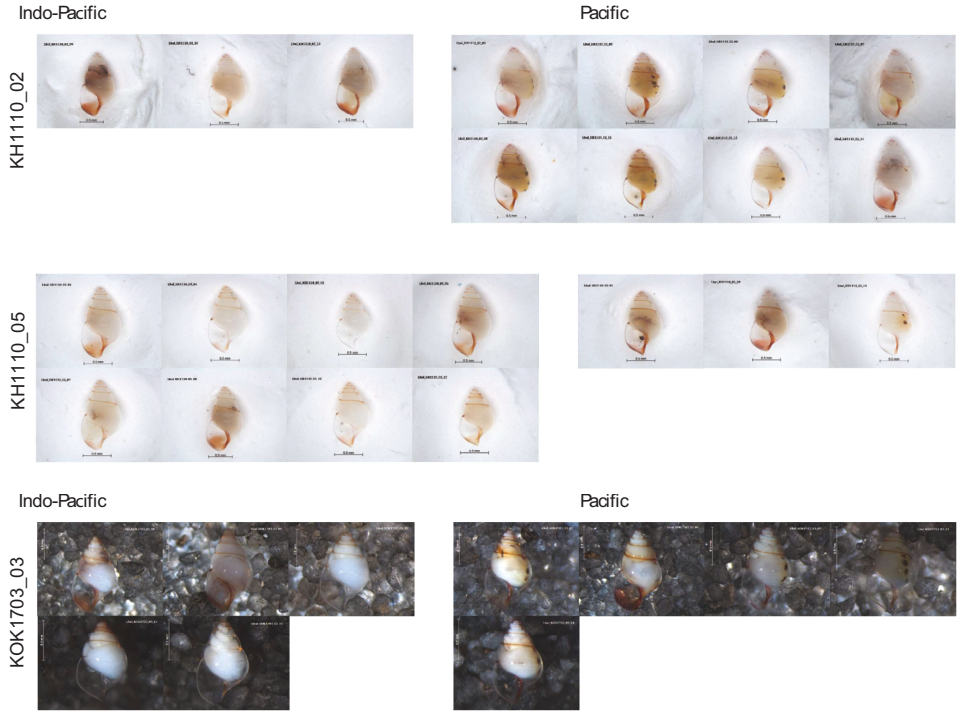


FIGURE S7 Shell images of all *Limacina bulimoides* included in this study. Shells are arranged according to sampling location and their genetic lineage. Specimens from the locations AMT24_17, AMT24_22, AMT24_23 and KOK1703_03 were photographed against a dark background, compared to the specimens from the other locations which were photographed against a light background. Higher resolution images are available on the Naturalis BioPortal database: https://bioportal.naturalis.nl/.

4.3.2. Colour assignment tables

TABLE S4 Aperture colour of each specimen of *Limacina bulimoides* per station, qualitatively assigned to the four colours: transparent, pink, tan and red-brown.

Station	Transparent	Pink	Tan	Red-brown
NIC2_S1C3	0	0	3	8
NIC8_S5C3	0	0	6	5
AMT24_17	0	11	0	0
NIC2_S9C3	0	0	0	10
AMT24_22	0	10	0	0
AMT24_23	4	3	0	3
SN105_08	0	0	0	10
KH1110_02	0	1	4	6
KH1110_05	1	2	4	4
KH1110_08	0	0	10	0
KH1110_15	4	7	0	0
KH1110_18	9	2	0	0
KH1110_21	10	1	0	0
KOK1703_03	1	4	0	5
SO255_143	0	11	0	0

TABLE S5 Aperture colour of each specimen of *Limacina bulimoides* per ocean basin, qualitatively assigned to the four colours: transparent, pink, tan and red-brown.

Location	Transparent	Pink	Tan	Red-brown
Atlantic	4	24	9	26
Indian	0	0	0	10
N. Pacific	2	7	18	15
S. Pacific	23	21	0	0

TABLE S6 Aperture colour of each specimen of *Limacina bulimoides* per lineage, qualitatively assigned to the four colours: transparent, pink, tan and red-brown.

Lineage	Transparent	Pink	Tan	Red-brown
Atlantic	4	24	9	26
Indo-Pacific	2	3	15	16
Pacific	23	25	3	9

4.3.3. Tissue pigmentation on *Limacina bulimoides*

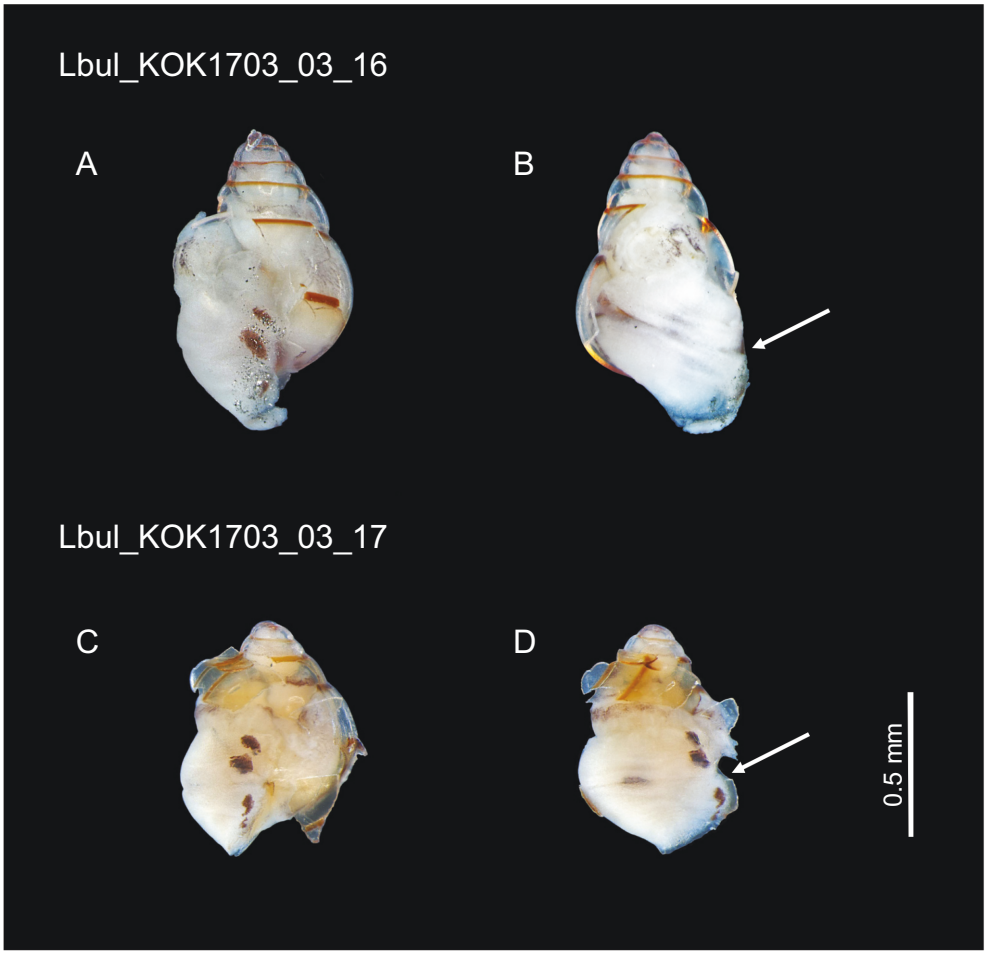


FIGURE S8 Images of two *Limacina bulimoides* individuals from the KOK1703_03 site that were not sequenced, but they possess the dark pigmented spots on their parapodia that are similar to those observed in the 16 Pacific lineage individuals (FIGURE S7). The shells were gently cracked to reveal the pigmentation on their tissue. (A, C) Four pigmented spots can be seen at the edges of the parapodia or 'wing feet'. There are two pigmented spots on each 'wing'. (B, D) The gap between the two parapodia is indicated by the white arrow.

5

Genome-based population structure of the pteropod *Limacina bulimoides* in the Atlantic Ocean

L.Q. Choo, G. Spagliardi, G. Hoarau, E. Goetze,
J. Huisman, K.T.C.A. Peijnenburg

Abstract

To date, there are many examples of population divergence in the open ocean, however, it is still unclear how divergence is maintained without obvious physical barriers between the populations. This is especially so for holoplanktonic organisms, which live in the water column for their entire lives and have high fecundities, leading to high potentials for gene flow and population connectivity. Knowledge of dispersal barriers in the open ocean can give us insight into speciation processes as well as information on how different populations and species can evolve in response to climate change. Previously, population structure was identified in the subtropical pteropod *Limacina bulimoides* collected from the Atlantic Ocean between 34°N-35°S, albeit with conflicting signals between the two barcoding genes used. To gain a clearer understanding of their population structure, we obtained genome-wide single-nucleotide polymorphisms across an Atlantic dataset of 142 *L. bulimoides* individuals collected in 2014 and 2017/2018, with the targeted enrichment of 2900 genomic regions. We observed that there was no recent gene flow between the North, Equatorial and South Atlantic populations, with narrow geographic barriers between these three populations at 14-15°N and 15-18°S. For six individuals, the mitochondrial COI haplotypes were incongruent with nuclear genetic variation, which is indicative of mitochondrial introgression during their evolutionary history, rather than the presence of expatriates which have traversed the dispersal barriers. The three populations were bounded by distinct ocean currents and productivity gradients, with the North and South populations found within the oligotrophic North and South Atlantic Gyre, while the Equatorial population was located within the more eutrophic boundaries of the North and South Equatorial Current. Additional demographic history and seascape genomics analyses could be conducted to further elucidate the drivers of this strong population structure in the Atlantic basin for *L. bulimoides*, and assess their ability to cope with future ocean changes.

Introduction

Genetic studies have uncovered unexpected population divergence in the open ocean (Bowen et al., 2016; Bucklin et al., 2018; Knowlton, 2000; Peijnenburg and Goetze, 2013). With no obvious barriers to dispersal and the constant movement of water masses, the dispersal potential of marine plankton populations in the open ocean was assumed to be unlimited (Norris, 2000; van der Spoel and Heyman, 1983). Geographically distant populations could still be connected by gene flow, especially in circumglobal species with long-lived pelagic durations or holoplanktonic species with large effective population sizes and high fecundity. While divergence by vicariance or geographic separation has been widely regarded as the most common form of speciation, examples of such allopatric modes of speciation from the open ocean are rare (but see Filatov et al., 2021). Most instances are recognised from well-known marine biogeographical barriers, such as the Isthmus of Panama (Knowlton et al., 1993), or well-characterised hybrid zones with opportunities for historical vicariance followed by secondary contact (Laakkonen et al., 2021; Simon et al., 2021). Yet, with the advent of genetic analyses, distinct populations and reproductive barriers have been identified in circumglobal species previously assumed to be homogeneous or panmictic (e.g., Addamo et al., 2020; Andrews et al., 2014; Cornils et al., 2017; Hirai et al., 2015; Pérez-Portela et al., 2013; Wall-Palmer et al., 2018). Their genetic divergence may be attributed to speciation in the presence of gene flow (Bierne et al., 2003; Johannesson, 2009, 2016; Peijnenburg and Goetze, 2013; Potkamp and Fransen, 2019) or the presence of dispersal barriers that have not yet been identified (Bowen et al., 2016; Faria et al., 2021), such as temperature, currents or a combination of other abiotic and biotic factors.

The potential drivers of genetic structuring in the open ocean can vary across scales. Currents can contribute to the dispersal of individuals and partitioning of populations from small spatial scales at tens of kilometres (White et al., 2010) to regional scales, like seas (Bertola et al., 2020; Hu et al., 2013) and ocean basin scales (Goetze, 2005; Richter et al., 2019; Xuereb et al., 2018a). Environmental gradients have also been shown to function as species-specific barriers to gene flow (Benestan et al., 2016; Bernatchez et al., 2019; Burridge et al., 2015; Saenz-Agudelo et al., 2015; Sandoval-Castillo et al., 2018) and drivers of general biogeographic patterns across taxa (Costello et al., 2017; Holman et al., 2021; Rombouts et al., 2009; Stanley et al., 2018). Within the Atlantic Ocean, the two oligotrophic subtropical gyres and the mesotrophic equatorial upwelling system form distinct habitats with characteristic environmental conditions, that have been identified as distinct 'ocean provinces' in the surface waters (Longhurst, 2007) and in the mesopelagic zone (Sutton et al., 2017). These gyres and the equatorial upwelling system are subject to seasonal and inter-annual environmental variation (Reygondeau et al., 2013), leading to changes in current direction, such as the presence of the east-flowing North Equatorial Current in the western tropical Atlantic only between the months July to September (Philander, 2001), and variation in the

location of the gyre boundaries over decadal scales (Drouin et al., 2021). Yet, for several holoplanktonic species in the Atlantic basin, spatial genetic structuring was found congruent with the gyre systems and environmental gradients (Goetze et al., 2015; Hirai et al., 2015; Kulagin et al., 2021; Norton and Goetze, 2013; Wall-Palmer et al., 2016a), and was shown to be temporally stable (Goetze et al., 2015).

Pteropods play important roles in the marine pelagic ecosystem. Pteropods are responsible for a biomass production of ~500 Tg of organic carbon per year across the global ocean (Bednaršek, Možina, Vogt, O'Brien, & Tarling, 2012), and contribute up to 89% of total calcification in pelagic waters (Buitenhuis et al., 2019). In addition, they form a significant part of the diet of fish and other predators in the pelagic food web (Groot and Margolis, 1991; Hunt et al., 2008; Lalli and Gilmer, 1989). Shelled pteropods possess thin aragonitic shells, which have been found to be vulnerable to dissolution. Pteropods show difficulties in maintaining shell thickness in present-day acidified waters (Busch et al., 2014; Mekkes et al., 2021a) and when exposed to predicted future conditions generated by ocean acidification scenarios (Bednaršek et al., 2018; Gardner et al., 2018; Manno et al., 2018; Mekkes et al., 2021b; Niemi et al., 2021). Understanding the spatial genetic structure of pteropod populations can allow us to detect evidence of local adaptation (Sanford and Kelly, 2011; Sgrò et al., 2011), and reconstruct the past evolutionary trajectories of species to understand their history of adaptation and speciation in marine environments (von der Heyden, 2017). With this knowledge, it is possible to gain insight into their adaptive potential and resilience to changing ocean conditions (Munday et al., 2013; Sanford and Kelly, 2011; Sgrò et al., 2011; Sunday et al., 2014)

Here, we investigate the population structure of the most abundant *Limacina* species in the tropical and subtropical ocean, *Limacina bulimoides*. We previously identified a primary dispersal barrier across a meridional Atlantic transect sampled in 2012 at 15-18°S based on shell shape, mitochondrial cytochrome oxidase I (COI) and nuclear ribosomal RNA 28S variation, as well as a secondary, more tentative barrier across the equatorial upwelling region supported only by 28S variation (Choo et al., 2021). In this study, we aim to characterise genome-wide variation in *L. bulimoides* within the Atlantic basin, with increased resolution based on additional sampling and by using 2900 previously designed genome-wide target capture probes from Choo et al. (2020). This allowed for a much more detailed insight into the population structure of this abundant pteropod in the Atlantic basin.

Materials and Methods

Sample collection

Plankton samples were collected in the Atlantic Ocean from two research expeditions, the Atlantic Meridional Transect (AMT) 24 cruise in 2014 and the Netherlands Initiative Changing Oceans (NICO) legs 2 and 8 in 2017 and 2018, respectively (Table 1, Figure 1). Specimens were collected via oblique tows to the surface using a 0.7 m

TABLE 1 Collection details of specimens of *Limacina bulimoides* from the Atlantic Ocean. Specimens were collected on Atlantic Meridional Transect cruise 24 (AMT24) in 2014 and during the Netherlands Initiative Changing Oceans (NICO) legs 2 and 8 in 2017 and 2018, respectively. The collection dates are in the format DD/MM/YYYY. Number of individuals used in genomic analyses using target capture are given in the last column.

No.	Latitude	Longitude	Station	Cruise	Max. sampling depth (m)	Date collected	Target capture
1	38°45′N	56°18′W	NIC8_S5	NICO-8	120	14/4/2018	11
2	35°58′N	69°50′W	NIC8_S4	NICO-8	125	10/4/2018	11
3	30°43′N	74°32′W	NIC8_S2	NICO-8	110	8/4/2018	10
4	27°30′N	28°53′W	AMT24_07	AMT24	328	5/10/2014	11
5	21°16′N	21°01′W	NIC2_S1	NICO-2	60	31/12/2017	11
6	20°27′N	29°16′W	AMT24_09	AMT24	274	7/10/2014	10
7	14°12′N	27°56′W	AMT24_11	AMT24	305	9/10/2014	9
8	5°58′N	47°50′W	NIC2_S9	NICO-2	80	13/1/2018	10
9	0°05′N	25°01′W	AMT24_15	AMT24	341	13/10/2014	9
10	7°28′S	25°07′W	AMT24_17	AMT24	266	15/10/2014	11
11	11°02′S	25°03′W	AMT24_18	AMT24	292	16/10/2014	10
12	18°19′S	25°05′W	AMT24_20	AMT24	283	18/10/2014	9
13	24°27′S	25°03′W	AMT24_22	AMT24	323	21/10/2014	10
14	27°46′S	25°01′W	AMT24_23	AMT24	260	22/10/2014	10

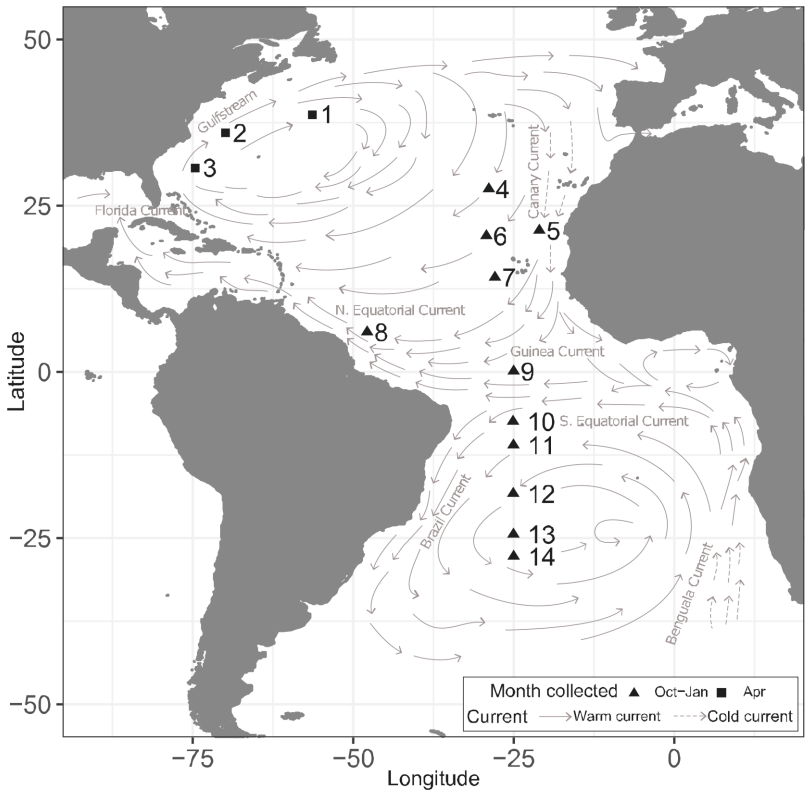


FIGURE 1 Sampling locations of *Limacina bulimoides* from the Atlantic Ocean. The sampling locations are labelled in order of descending latitude, according to TABLE 1.

diameter bongo net with a mesh size of 200 µm or a 1 m diameter ring net with a mesh size of 350 µm, with a tow duration of 14 to 50 minutes and a maximum depth range between 60 and 341 m. Samples were immediately preserved on board in 96% ethanol and stored at -20 °C, with replacement of the ethanol after 24 h of preservation. *Limacina bulimoides* specimens were subsequently sorted in the laboratory and photographed prior to destructive genetic work.

Library prep, sequencing and quality filtering

Genomic DNA was separately extracted from each of the 142 individuals with either the E.Z.N.A mollusc or insect kit (Omega Bio-Tek). The DNA was sheared by sonication to attain a peak length of 300bp with a Covaris S220 or ME220 focused ultrasonicator. After sonication, the fragmented DNA was prepared into individual libraries using the NEXTflex Rapid Pre-Capture Combo Kit (Bioo Scientific). Individually barcoded libraries were subsequently pooled at equimolar concentrations with 26-27 libraries per pool. The target capture reaction was then performed on each pool using the myBaits Custom Target Capture kit (Arbor Biosciences), with a capture probe set specifically designed for *L. bulimoides*, as described in Choo et al. (2020). Captured library pools were sequenced on the Illumina NextSeq 500 platform using high-output v2 chips (150 cycles).

Raw sequences were de-multiplexed (Genbank accessions: SAMN11131474-82, SAMN20293115-247, SAMN21503552-627) and then mapped with BWA 0.7.12 (Li, 2013) to a reduced contig set of the genomic assembly (Genbank accession: SWLX00000000). The resulting alignments were cleaned and filtered with SAMtools version 1.4.1 (Li et al., 2009) to retain only properly and uniquely mapped paired reads. Duplicates were marked and removed with Picard version 2.18.5 (http://broadinstitute.github.io/picard). Variant calling was done using GATK 4.1.7.0, following the Variant Discovery Pipeline (Auwera et al., 2013; Depristo et al., 2011) with GNU parallel utility (Tange, 2011) for executing commands in parallel. Haplotypes were called individually using HaplotypeCaller with emitRefConfidence output with and without the setting EMIT_ALL_CONFIDENT_SITES in two separate instances. The resulting gVCF files for each setting were combined with CombineGVCFs. The combined gVCF files were then genotyped using GenotypeGVCFs. Single nucleotide polymorphisms (SNP) were then extracted from the total variants using SelectVariants (-SelectType SNP). SNPs were hard-filtered with VariantFiltration using QualByDepth (QD) <2.0, FisherStrand (FS) >60.0, RMSMappingQuality <5.0, MQRankSumTest (MQRankSum) <-5.0, ReadPositionRankSum (ReadPosRankSum) <-5.0 to retain reliable SNPs. The SNPs were subsequently processed in BCFtools version 1.7 (https://github.com/samtools/bcftools) to retain only high-quality SNPs (Phred score >20) that had a coverage of at least three.

Population structure

We assessed population structure using the filtered SNP dataset, with additional filters to retain only SNPs present in at least 80% of the genotypes, with a minor

allele frequency (MAF) <0.01 and without excess heterozygosity that can indicate badly mapped regions. From those filtering settings, we retained 223,058 nuclear SNPs, after removing the three SNPs that were mapped to two known contigs of mitochondrial origin. This was done to avoid conflicting signals between the nuclear and mitochondrial SNPs, which could have distinct evolutionary histories. To visualise population structure, we conducted a Principal Component Analysis (PCA) in PLINK v1.90b6.17 (Chang et al., 2015), using linkage pruning settings of window size of 50 variant counts, window shift of 10 variant counts and r^2 of 0.2 to retain 97,425 SNPs in linkage equilibrium. In addition, we used Admixture (Alexander et al., 2009) to infer populations and ancestries, with the same linkage pruned SNP dataset for PLINK. The analysis was conducted for values of K between 1 and 7, and the cross validation error for each K value was calculated to determine the number of putative ancestral populations.

In addition to clustering analyses, we investigated population structure based on nearest neighbour haplotype co-ancestry with fineRADstructure v.0.3.2 (Malinsky et al., 2018). We used the hapsFromVCF function of RADpainter to convert the filtered nuclear SNP dataset used above, without linkage pruning, into input format with SNPs from each of the 2641 contigs condensed into haplotypes. A locus was considered missing if more than half of the SNPs were missing. Of the 142 individuals, two were excluded as they had a higher proportion (>4%) of loci which were marked as missing. After excluding these individuals, we used the paint function of RADpainter to calculate a co-ancestry matrix, which summarises the nearest neighbour haplotype relationships in the data set. Next, a clustering dendrogram of shared ancestry was inferred from the co-ancestry matrix using the fineSTRUCTURE Markov chain Monte Carlo clustering algorithm, with 100,000 burn-in iterations, 100,000 sample iterations, and thinning of 1000. The inferred clusters were arranged with a simple tree building algorithm in fineSTRUCTURE with 10,000 hill-climbing iterations.

Genetic signals from the targeted mitochondrial COI fragments were separately analyzed in a haplotype network after mapping and *de novo* assembly in Geneious Prime 2021.1.1 (https://www.geneious.com). First, all raw Illumina NextSeq reads of each of the 142 individuals were mapped to a reference database of *L. bulimoides* COI sequences (NCBI: MN952611-MN952965), using the Geneious assembler with medium-low sensitivity. The mapped reads were extracted and *de novo* assembled per individual using the Geneious assembler with medium sensitivity. The resulting contigs were annotated for the COI region, and the COI annotation (564 bp) with the highest coverage was extracted for each individual, and subsequently translated (Translation TABLE 5) to check that no stop codons or gaps were present. The 142 COI annotations (Genbank accessions: MZ542566-596, MZ542673-704, OK185211-289) were combined with one representative each of the known Atlantic haplogroups 1 and 2 from Choo et al. (2021) (Genbank accessions: MN952611, MN952937), and used to create a multiple sequence alignment

using MAFFT v7.222 (Katoh and Standley, 2013). A minimum spanning network was calculated from the alignment and visualised in POPART (Leigh and Bryant, 2015).

Diversity indices

To estimate genetic diversity and the extent of genetic structuring within populations, we calculated heterozygosity, nucleotide diversity, global weighted F_{ST} and pairwise F_{ST}. Heterozygosity and nucleotide diversity were calculated with the same filters of retaining only SNPs present in at least 80% of the genotypes, and removing sites with excess heterozygosity for vcf files containing SNPs (n = 661,912) and all confidently called variant sites (n = 788,818), respectively. Global weighted F_{ST} and pairwise F_{ST} were calculated from filtered vcf files above with an additional filter for MAF < 0.01 with 223,061SNPs.

Heterozygosity was calculated in VCFtools 0.1.15 (Danecek et al., 2011) with the setting –het, which outputs a measure of heterozygosity on a per-individual basis. We plotted these values per population to obtain the population-specific distribution of heterozygosity. Nucleotide diversity was calculated in VCFtools 0.1.15 (Danecek et al., 2011) for each population, with the SNPs partitioned into their respective populations and a window size of 2865 bp corresponding to the average contig size. Distributions of heterozygosity and nucleotide diversity for each population were plotted in R v.4.0.3 (R Core Team, 2017) and a one-way ANOVA was used to test for significant differences between populations. Global weighted F_{ST} for each population was calculated with the MAF-filtered SNP dataset for each population, by having one SNP randomly chosen per contig according to a custom script (Choquet et al., 2019), and calculating the weighted F_{ST} in PLINK v1.90b6.17 (Chang et al., 2015) for each thinning iteration for 1000 iterations. The resulting distributions for each population were plotted in R (R Core Team, 2017). Pairwise F_{ST} was similarly calculated with one SNP per contig in VCFtools 0.1.15 (Danecek et al., 2011), with the –weir-fst-pop setting and averaged over 1000 iterations for each pairwise comparison of sampling stations within populations.

Results

Data processing

For the 142 *L. bulimoides* individuals sampled, 10,564,661 mean raw reads per individual were sequenced. To retain only confidently called SNPs, we filtered the dataset for reads that could be mapped to the reduced genome assembly with high quality and in proper pairs. An average of 4,228,701 reads per individual could be mapped to the reduced genomic assembly (Table S1). Of these reads, 40.9% were mapped with high quality and 93.3% of high quality reads were on target, indicating that the remaining 6.7% most likely mapped to regions in the contig adjacent to the target. The total number of confidently called SNPs obtained for the 142 individuals before and after hard filtering were 1,776,923 and 1,434,869, respectively,

while the average sequencing depth across all individuals after hard filtering was 288x. Quality filtering was used to retain sites with a depth of more than or equal to three and a minimum genotype quality of 20. Downstream filtering settings were adjusted according to the type of analysis to be performed (Figure S1).

Genome-wide population structure

Within the Atlantic basin, we found evidence supporting three geographically separate populations, located in the North, Equatorial and South Atlantic (Figure 2). The North Atlantic population was comprised of 64 individuals from six sampling locations distributed along the Gulf Stream and the Canary Current. The Equatorial Atlantic population consisted of 49 individuals from five locations in the Atlantic equatorial upwelling region, while the South Atlantic population was comprised of 29 individuals from three locations in the South Atlantic Gyre. Based on the Principal Component Analysis, no intermediates or expatriate individuals were observed between any of the populations (Figure 2B). On PC1, which comprised 14% of total genetic variation, the Equatorial population was mainly separated from the North and South Atlantic population, while PC2, which comprised 12.4% of total genetic variation, separated the South population from the North and Equatorial populations. The same three populations were also recovered in the fineRADstructure co-ancestry matrix showing the absence of any mixing of individuals between populations (Figure 3). The largest amount of genetic divergence was between the Equatorial and South Atlantic (F_{ST} = 0.197), followed by North and South Atlantic (F_{ST} = 0.126) and North and Equatorial Atlantic (F_{ST} = 0.122).

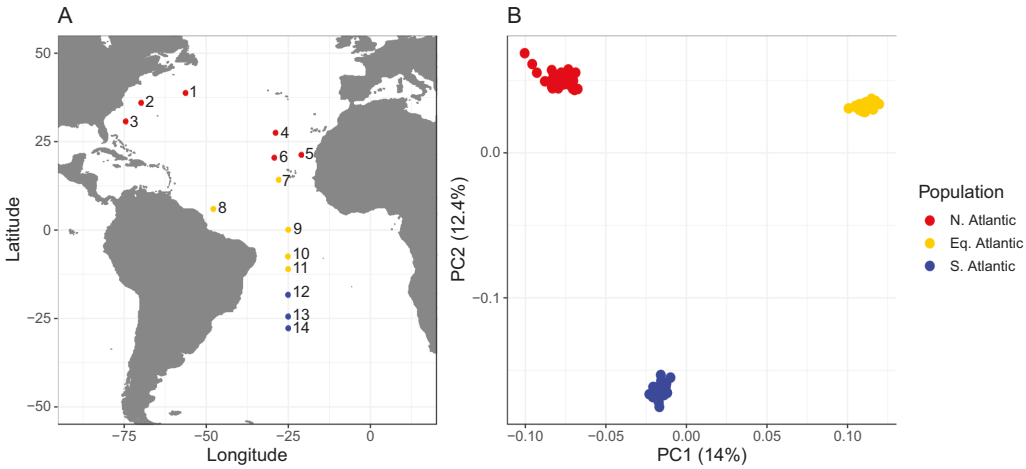


Figure 2 (A) Location of the 14 sampling sites of *Limacina bulimoides* labelled according to Table 1 and coloured according to their genome-wide population structure, which revealed distinct populations in the North Atlantic (red), Equatorial Atlantic (yellow) and South Atlantic (blue). (B) Principal Component Analysis (PCA) plot based on 97,425 nuclear single nucleotide polymorphisms across142 individuals, coloured according to their population structure.

Mitochondrial population structure

The mitochondrial COI fragment was characterised by high variability, with all but two individuals having unique haplotypes (Figure 4). The two individuals with the

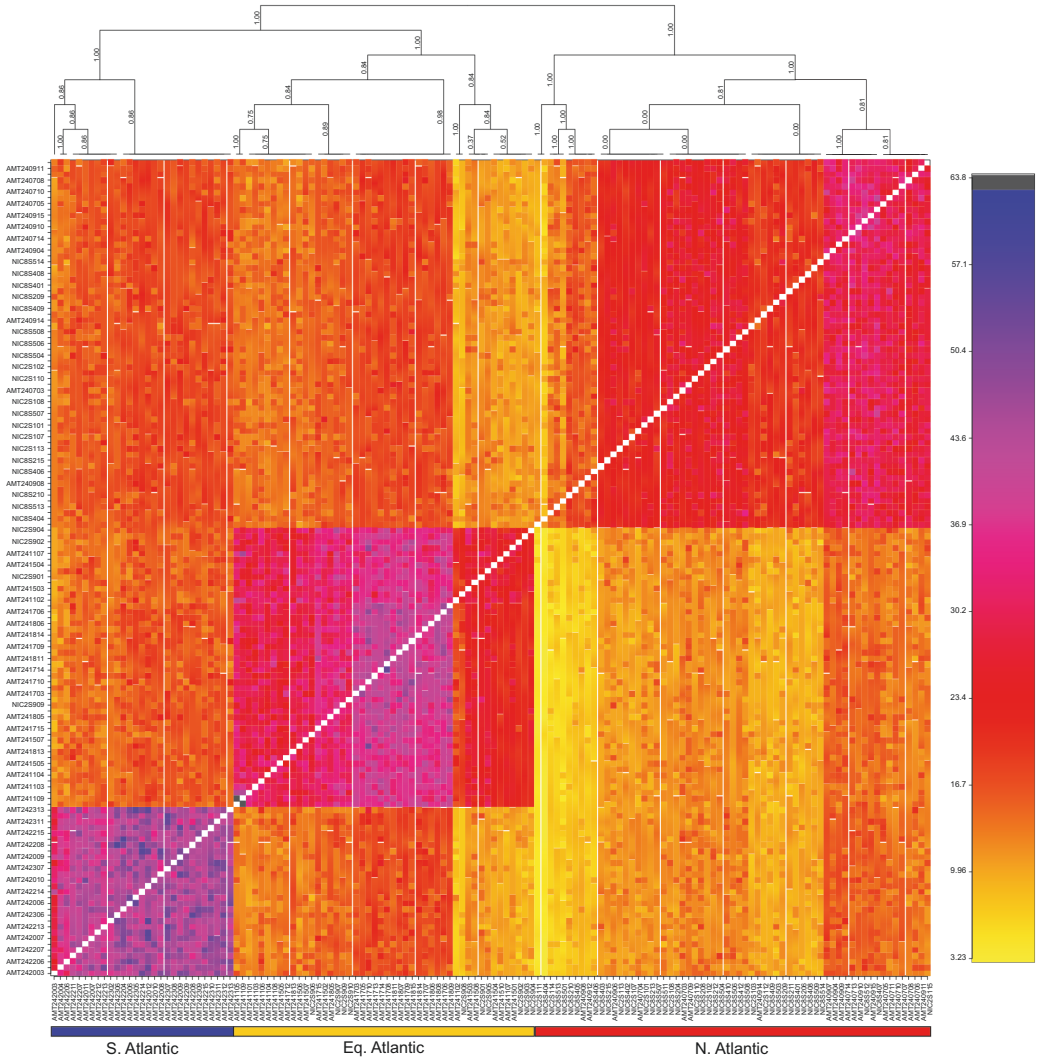


Figure 3 Co-ancestry matrix generated using fineRADstructure, which clusters the 142 individuals of *Limacina bulimoides* according to their nearest neighbour relatedness. Plot is coloured according to the co-ancestry value between pairs of individuals (see legend: black/blue colours indicate more relatedness while yellow indicates less relatedness). The three clusters identified correspond with the three geographically-distributed populations in the North, Equatorial and South Atlantic (see also Figure 1). Note the presence of two highly related individuals (two black squares adjacent to diagonal line) within the Equatorial Atlantic cluster, identified as Lbul_AMT24_11_05 and Lbul_AMT24_11_09.

same COI haplotype were from the North Atlantic, with one being the reference Haplogroup 1 sequence from Choo et al. (2021). Two main haplogroups separated by at least 17 nucleotide substitutions were recovered; they corresponded to Haplogroup 1 (composed of mostly North and Equatorial Atlantic individuals) and Haplogroup 2 (composed of South Atlantic individuals) previously identified in a different dataset sampled in 2012 (Choo et al., 2021). Six individuals possessed a COI haplotype that grouped with Haplogroup 1 (North and Equatorial Atlantic) while originating from the South Atlantic (blue dots in Figure 4). These individuals were mostly located on long branches and on the outside of the network.

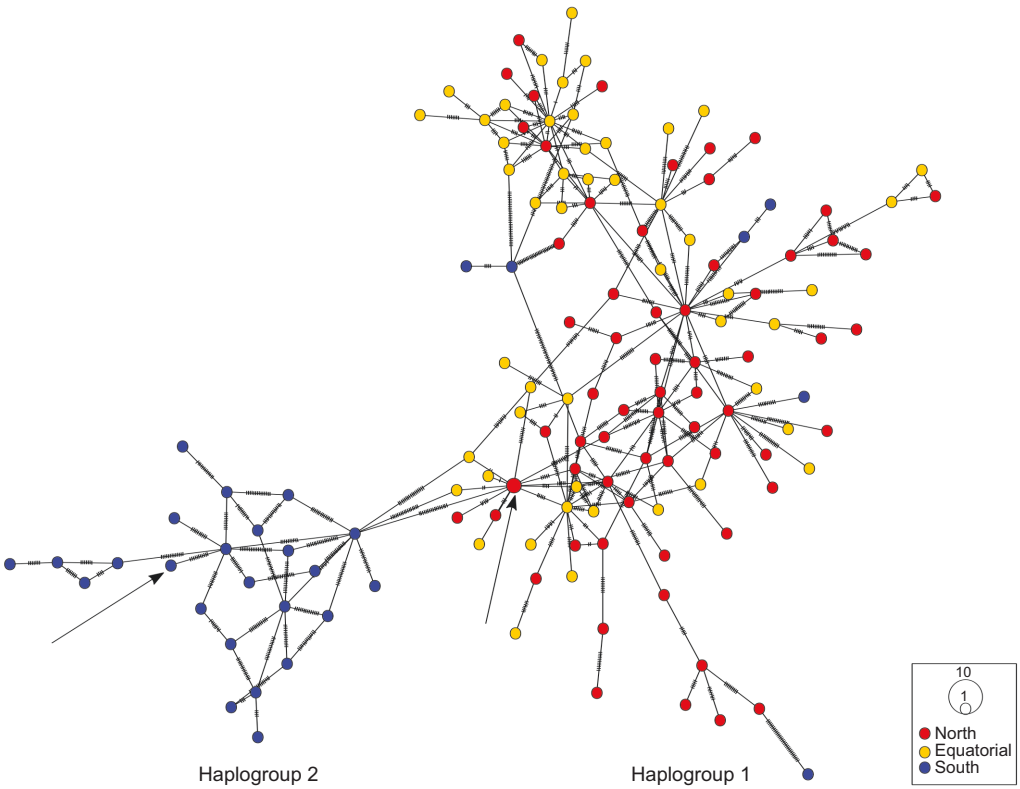


Figure 4 Minimum spanning network of the mitochondrial cytochrome oxidase I (COI) gene fragment. The size of the dots represents the number of individuals with each haplotype, with most dots representing one individual. The colour of the dots indicates the populations identified in the North Atlantic (red), Equatorial Atlantic (yellow) and South Atlantic (blue). Two representative sequences from the previously documented Haplogroup 1 and 2 in the Atlantic that were sequenced in Choo et al. (2021) are included in this network; they are indicated with an arrow pointing to these haplotypes (Genbank accessions: MN952611, MN952937). Haplogroup 1 was located in the North and Equatorial Atlantic, while Haplogroup 2 was located in the South Atlantic.

TABLE 2 Pairwise F_{ST} comparisons across sampling stations within each of the three populations of *Limacina bulimoides*: North, Equatorial and South Atlantic. Pairwise F_{ST} values were averaged over 1000 iterations for each pairwise comparison, with each iteration including one randomly selected single nucleotide polymorphism per contig.

North	1	2	3	4	5
1. NIC8_S5					
2. NIC8_S4	0.00163				
3.NIC8_S2	0.00015	-9.77E-07			
4. AMT24_07	0.00191	0.00193	0.00164		
5. NIC2_S1	0.00121	0.0024	-0.00029	0.00161	
6. AMT24_09	0.00218	0.00347	0.00303	0.00119	0.00217
Equatorial	7	8	9	10	
7. AMT24_11					
8. NIC2_S9	0.00265				
9. AMT24_15	0.00173	-0.0009			
10. AMT24_17	0.0034	0.00146	0.00203		
11. AMT24_18	0.00316	0.00097	0.00188	-0.00073	
South	12	13			
12. AMT24_20					
13. AMT24_22	-0.00139				
14. AMT24_23	0.00145	-0.000356			

VARIABILITY WITHIN POPULATIONS

Within-population structure was relatively low compared to between-population comparisons, based on pairwise F_{ST} values (within-population comparison range: -0.00139 to 0.00347; between-population comparison range: 0.122-0.197), indicating a high level of shared genetic variation within populations. The South Atlantic seemed to be relatively homogeneous across the three sampling locations, with relatively low pairwise F_{ST} comparisons between stations (TABLE 2) and the highest amount of within-population relatedness given by the co-ancestry matrix (FIGURE 3). However, this could be an artefact of the limited geographic sampling for the South population as compared to the North and Equatorial populations. The North and Equatorial populations had lower relatedness overall and higher variability in relatedness between individuals (FIGURE 3). Within the North Atlantic population, individuals from station AMT24_09 (nr. 6) had a larger genetic distance to individuals from the other stations in the population; while within the Equatorial population, individuals from station AMT24_11 (nr. 7) had a larger genetic distance to those from the other stations in the population (TABLE 2). Interestingly, two closely-related individuals were detected in the Equatorial population based on their high amount of pairwise co-ancestry, as indicated by the black boxes adjacent to the diagonal line (FIGURE 3).

GENETIC DIVERSITY

Mean heterozygosity was not significantly different between the North Atlantic, Equatorial and South Atlantic population (FIGURE 5A; One-way ANOVA: $F_{(2,139)} = 0.914$, $p = 0.403$). The South Atlantic population had the highest nucleotide diversity

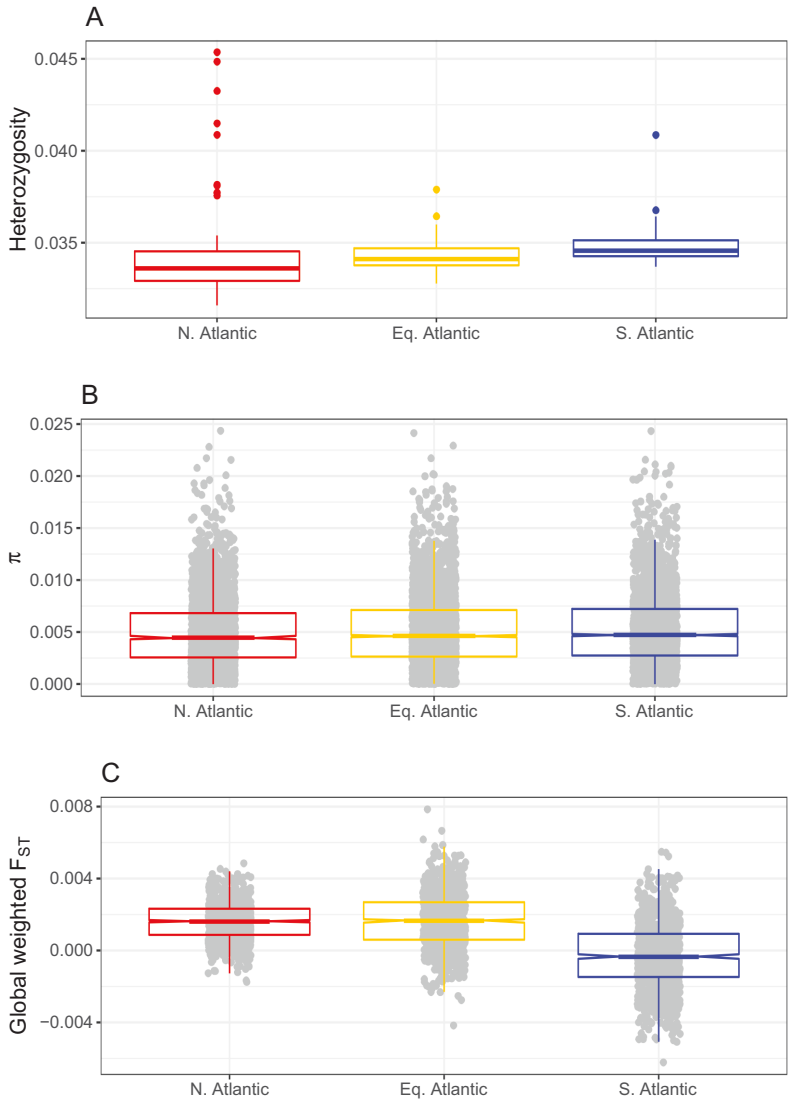


FIGURE 5 Distribution of heterozygosity, nucleotide diversity (π) and global weighted F_{ST} for each of the three populations of *Limacina bulimoides* distributed in the North (N = 64), Equatorial (N = 49) and South (N= 29) Atlantic. The analysis is based on filtered single nucleotide polymorphism (SNP) datasets for heterozygosity and global weighted F_{ST}, while a dataset of all confidently called sites was used to calculate nucleotide diversity. The lower and upper hinges of the boxplots represent the first and third quartile and span the interquartile range (IQR), while the median is indicated by the horizontal line between the hinges of the boxplot. (A) Individual heterozygosity. (B) Nucleotide diversity (π) of SNPs in each population calculated across a window size of 2865 base pairs. (C) Global weighted F_{ST} calculated across 1000 iterations for each population, with each iteration including one randomly chosen single nucleotide polymorphism per contig.

(π = 0.00529), followed by the Equatorial Atlantic (π = 0.00520) and North Atlantic (π = 0.00503) populations (Figure 5B), with a significant difference (one-way ANOVA: $F_{(2,8563)}$= 3.983, p = 0.0187) in nucleotide diversity between the North and South Atlantic (Tukey HSD: North – South: p-adjusted = 0.0153, North – Equatorial: p-adjusted = 0.157, Equatorial – South: p-adjusted = 0.615). Global weighted F_{ST} was significantly larger than zero for both the North Atlantic (mean = 0.00161, t = 49.398, df = 999, p < 2.2e-16) and Equatorial Atlantic (mean = 0.00167, t = 34.053, df = 999, p < 2.2e-16), indicating the presence of population substructure. For the South Atlantic population, the distribution of global weighted F_{ST} was not significantly different from zero (mean = -0.000323, t = -5.51, df = 999, p = 1) (Figure 5C, see also Table 2).

Discussion

We identified three distinct populations of *Limacina bulimoides* within the Atlantic basin, with no evidence of mixing between populations despite the continuous pelagic environment. The three populations identified, the North, Equatorial and South Atlantic (Figure 2), correspond with the populations identified across two dispersal barriers based on the 28S gene in a previous barcoding study (Choo et al., 2021). The consistent population structure in the Atlantic from 2012 (Choo et al., 2021) to 2014, 2017 and 2018 (present study) indicates the temporal persistence of this pattern of genetic partitioning.

Migration across the two dispersal barriers is unlikely to have occurred in recent times, instead, the cases of incongruence observed between the mitochondrial haplogroups and their sampling location are probably the result of introgression. Genome-wide variation showed clearly that there are no hybrids between populations (Figure 2). However, the COI gene paints a conflicting picture, with the presence of six individuals that have a conflicting mitochondrial haplogroup compared to their nuclear genomic variation which is congruent with their sampling location (Figure 4). We previously suggested from a dataset containing only COI and 28S barcoding genes that such individuals represent expatriates that had dispersed across the population barrier (Choo et al., 2021). However, with the added perspective from genome-wide SNPs, it is more likely that the six individuals (Figure 4) are not recent migrants or expatriates. Rather, the mismatch between their COI haplogroup and nuclear genomic variation represents the legacy of mitochondrial introgression during their evolutionary history. These discordant mitochondrial haplogroups seem to occur at a roughly consistent rate across our two datasets, with 3.65% discordant individuals in 2012 (Choo et al., 2021) and 4.22% in the present dataset consisting of individuals sampled across 2014 and 2017/2018. While we observed only introgression of Haplogroup 1 (North and Equatorial) into South Atlantic individuals in the current dataset of 142 individuals, we detected bidirectional introgression between Haplogroups 1 and 2 in our previous, larger dataset of 356 individuals in Choo et al. (2021). Hence it is likely that the introgression occurs in both ways across the dispersal barrier. Examples of bidirectional movement of mtDNA are less com-

mon than instances of asymmetric introgression (Sloan et al., 2017; Toews and Brelsford, 2012). The few documented instances of bidirectional introgression have been attributed to mtDNA heteroplasmy and paternal leakage (Mastrantonio et al., 2019) or the presence of weaker barriers to introgression in mtDNA compared to nuclear DNA (Colliard et al., 2010). A possible scenario for bidirectional mitochondrial introgression may be historical isolation followed by secondary contact (Toews and Brelsford, 2012) between genetically distinct and hybridising populations (North + Equatorial and South) across the dispersal barrier in the South Atlantic subtropical gyre. This may have occurred during the climatic oscillations associated with glacial-interglacial periods in the late Pleistocene, resulting in range shifts, temporary separation and subsequent re-mixing of populations (Hewitt, 2011). Alternatively, mitochondrial introgression could also occur in the absence of geographical isolation, as a consequence of selective advantages associated with environmental characteristics (Toews and Brelsford, 2012). We would require additional and more in-depth studies into introgressed individuals to determine the evolutionary mechanisms which resulted in this pattern of discordance. Here, we emphasise the importance of using genome-wide markers in addition to mitochondrial barcoding genes for understanding the complete evolutionary history of *L. bulimoides*, given their mitonuclear discordance (Rubinoff and Holland, 2005).

The dispersal barriers separating the three populations could be attributed to ocean currents separating the populations, ecological selection across the oceanic environmental clines, or both of these factors acting in concert. The geographic boundaries of the three populations are broadly demarcated by the major ocean currents (Figure 1) as well as transitions in ocean primary production (represented by the chlorophyll *a* concentration; Figure 6). The North Atlantic population is bounded by the Gulf Stream and North Atlantic Gyre, and characterised by oligotrophic surface conditions, while the South Atlantic population is located in the oligotrophic South Atlantic Gyre and extending into the Subtropical Convergence Zone with colder waters and more eutrophic conditions. The Equatorial Atlantic population occupies slightly more eutrophic, warm waters in the equatorial region, including both the North Equatorial Current and the South Equatorial Current. Additional sampling at the boundary regions and connected current systems, such as the Florida and Brazil currents, as well as comprehensive genome-wide genotyping of both juveniles and adults across the dispersal barriers will be needed to accurately determine the spatial ranges of these Atlantic populations and assess whether these dispersal barriers are indeed non-permeable.

When taken together with the location of dispersal barriers identified in a previous latitudinal dataset of *L. bulimoides*, and assuming that the locations of the barriers are stable across the years from autumn 2012 and autumn 2014, we find that the northern Atlantic subtropical barrier is located within a narrow region at 14-15°N and the southern Atlantic subtropical barrier is located between 15-18°S along the transects sampled. The presence of narrow barriers between the three distinct

populations suggests that isolation-by-distance is not the primary driver behind Atlantic population structure of *L. bulimoides*. Instead the narrow latitudinal barriers we observed may be due to historical isolation following glacial-interglacial cycles in the Pleistocene, followed by secondary contact of the reproductively isolated populations (Filatov et al., 2021), or ecological selection, where barriers represent regions of suboptimal habitat within which there are insufficient individuals for maintenance of a stable population (Goetze et al., 2017). In other holoplanktonic species, physical retention in gyre systems could play a role in population structuring as they show distinct boundaries linked to gyre systems (Andrews et al., 2014; Goetze et al., 2015, 2017) and North and South Atlantic subtropical gyral populations separated by an equatorial dispersal barrier (Norton and Goetze, 2013). The northern Atlantic subtropical barrier for *L. bulimoides* coincides with a mesopelagic boundary at ~14-22°N, identified in three species of Atlantic krill (Kulagin et al., 2021) and the northernmost population boundary for a mesopelagic copepod (Goetze et al., 2017), which suggests the presence of a hydrological barrier or phylogeographic break. The south Atlantic subtropical barrier at 15-18°S is also identified as a species boundary between two mesopelagic shrimp species (Judkins, 2014) and a mesopelagic boundary between two ecoregions (Sutton et al., 2017). Both boundaries seem to affect other pelagic taxa in addition to *L. bulimoides*, which may indicate the contribution of subtle shifts in multiple environmental factors as a taxa-

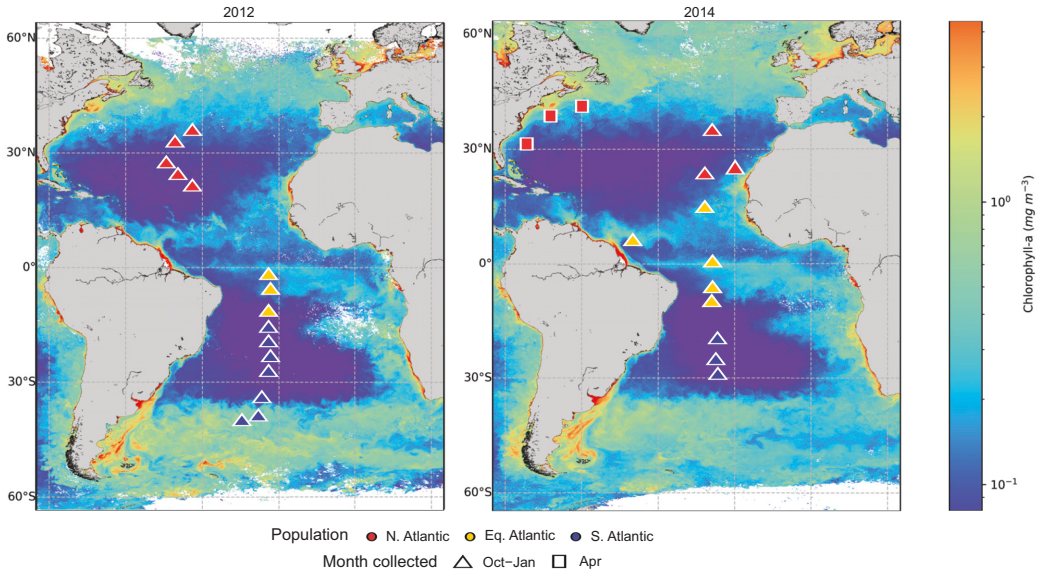


Figure 6 Spatial genetic structure of three populations of *Limacina bulimoides* in the Atlantic basin (North Atlantic: red, Equatorial Atlantic: yellow, South Atlantic: blue) based on the 2012 data from Choo et al. (2021) and the 2014 data of the present study superimposed on composite satellite images of chlorophyll *a* concentrations during the AMT22 (Oct-Dec 2012) and AMT24 (Oct-Dec 2014) cruises, respectively.

wide structuring force for divergence in the open ocean (Kulagin et al., 2021). Environmental conditions are known to fluctuate seasonally and inter-annually in the tropical and subtropical Atlantic leading to changes in levels of primary productivity (Barton et al., 2015; Franco et al., 2020; Pastor et al., 2013; Pérez et al., 2005), however, projected climate-driven oceanographic changes may still impact the stability of the population boundaries in *L. bulimoides* (Wilson et al., 2016).

Within the three populations, subtle genetic variability indicates that despite the high dispersal potential of *L. bulimoides* in a pelagic environment, populations still show some degree of further structuring. The slight genetic differences (Table 2) could indicate either small scale genetic patchiness or the influence of dispersal due to mixing of ocean currents and temporal effects. Genetic structure was detected within the North and Equatorial Atlantic populations but not in the South Atlantic population (Figure 2, 4C), which is likely to be a result of the limited geographic area sampled for the South population. In the North Atlantic population, higher F_{ST} values (Table 2) were found between the AMT sites, sampled in October 2014, and the NICO sites, sampled in December 2017/April 2018, regardless of their physical proximity. Hence, genetic variability within the three populations may be more related to seasonal or inter-annual variation rather than isolation-by-distance effects. Within the Equatorial Atlantic population, specimens from the boundary location AMT24_11 (nr. 7), were more genetically distinct from the other sampling locations further south (Table 2). Repeated sampling across environmental clines and different seasons will be needed to elucidate the fine-scale drivers affecting variability within populations.

In conclusion, there are three distinct populations of *L. bulimoides* within the Atlantic basin, distributed across the two subtropical gyres and the equatorial region. The absence of hybrids indicates that there has not been any recent gene flow between these populations, although the mitochondrial introgression detected across the southern barrier hints at admixture during their evolutionary history. This genome-wide study lends itself to further analyses including demographic inferences to determine the order and timing of divergence between North, Equatorial and South Atlantic populations, as well as their population size variation across time. In particular, seascape analyses combining genetic data, environmental variables and oceanographic distances could provide further insight into the drivers of population structure in the Atlantic Ocean.

Acknowledgements

We thank L. Mekkes, A. Burridge and M. Jungbluth for assistance at sea, and all captains and crews of the ocean expeditions for their support and assistance. We would also like to thank Dan Clewley and the Natural Environment Research Council Earth Observation Data Acquisition and Analysis Service (NEODAAS) for producing Figure 6. This research was supported by the Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to K.T.C.A.P. The Netherlands Initiative Changing Oceans (NICO) expedition on *RV Pelagia* was also funded by NWO and the Royal Netherlands Institute for Sea Research (NIOZ). The Atlantic

Meridional Transect is funded by the UK Natural Environment Research Council through its National Capability Long-term Single Centre Science Programme, Climate Linked Atlantic Sector Science (grant number NE/R015953/1). This study contributes to the international IMBeR project and is contribution number 374 of the AMT programme.

SUPPLEMENTARY MATERIAL

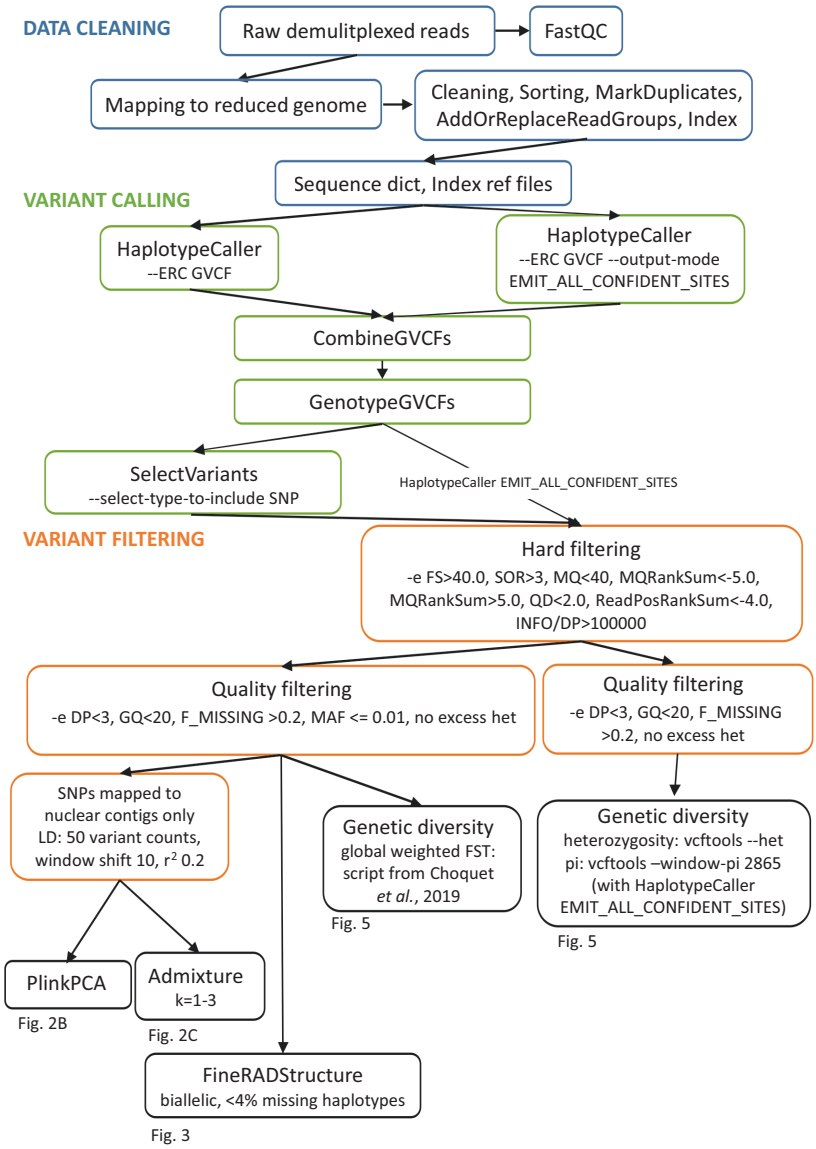


FIGURE S1 Single nucleotide polymorphism (SNP) calling pipeline and settings for downstream genetic analyses.

Table S1 Sampling location, assigned population, station name, museum voucher code, Biosample and mitochondrial cytochrome oxidase I (COI) accessions and capture efficiency statistics for each individual.

Specimen ID	Population	Station	Museum voucher code	Biosample	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth
Lbul_AMT24_07_03	N. Atlantic	AMT24_07	RMNH.MOL.340294	SAMN21503552	OK185211	5248097	1568316	29.9	87.5	97.38
Lbul_AMT24_07_04	N. Atlantic	AMT24_07	RMNH.MOL.340295	SAMN21503553	OK185212	4519347	1192130	26.4	88.2	74.67
Lbul_AMT24_07_05	N. Atlantic	AMT24_07	RMNH.MOL.340296	SAMN21503554	OK185213	9009899	3306148	36.7	90.2	214.11
Lbul_AMT24_07_06	N. Atlantic	AMT24_07	RMNH.MOL.340297	SAMN21503555	OK185214	8934328	3734474	41.8	91.4	245.98
Lbul_AMT24_07_07	N. Atlantic	AMT24_07	RMNH.MOL.340298	SAMN21503556	OK185215	8845084	4864454	55.0	93.7	332.98
Lbul_AMT24_07_08	N. Atlantic	AMT24_07	RMNH.MOL.340299	SAMN21503557	OK185216	6300860	3495096	55.5	89.3	226.54
Lbul_AMT24_07_09	N. Atlantic	AMT24_07	RMNH.MOL.340300	SAMN21503558	OK185217	8582723	4315704	50.3	88	273.56
Lbul_AMT24_07_10	N. Atlantic	AMT24_07	RMNH.MOL.340301	SAMN21503559	OK185218	9023927	4664836	51.7	87.4	293.18
Lbul_AMT24_07_11	N. Atlantic	AMT24_07	RMNH.MOL.340302	SAMN21503560	OK185219	10607071	4838244	45.6	92.3	324.13
Lbul_AMT24_07_13	N. Atlantic	AMT24_07	RMNH.MOL.340303	SAMN21503561	OK185220	8132049	4213356	51.8	92.7	284.18
Lbul_AMT24_07_14	N. Atlantic	AMT24_07	RMNH.MOL.340304	SAMN21503562	OK185221	14249089	5679000	39.9	89.7	365.42
Lbul_AMT24_09_04	N. Atlantic	AMT24_09	RMNH.MOL.340305	SAMN21503563	OK185222	9263744	2268452	24.5	85.7	136.88
Lbul_AMT24_09_06	N. Atlantic	AMT24_09	RMNH.MOL.340306	SAMN21503564	OK185223	3380927	1456102	43.1	92.9	97.93
Lbul_AMT24_09_07	N. Atlantic	AMT24_09	RMNH.MOL.340307	SAMN21503565	OK185224	12136250	5737698	47.3	90.9	375.28
Lbul_AMT24_09_08	N. Atlantic	AMT24_09	RMNH.MOL.340308	SAMN21503566	OK185225	5870716	3430796	58.4	93.9	235.35
Lbul_AMT24_09_09	N. Atlantic	AMT24_09	RMNH.MOL.340309	SAMN21503567	OK185226	10601848	6037416	56.9	89.2	389.8
Lbul_AMT24_09_10	N. Atlantic	AMT24_09	RMNH.MOL.340310	SAMN21503568	OK185227	13337321	7284048	54.6	88.3	463.83
Lbul_AMT24_09_11	N. Atlantic	AMT24_09	RMNH.MOL.340311	SAMN21503569	OK185228	8748775	4594890	52.5	87.7	289.5
Lbul_AMT24_09_12	N. Atlantic	AMT24_09	RMNH.MOL.340312	SAMN21503570	OK185229	8305782	4351124	52.4	92.7	294.13
Lbul_AMT24_09_14	N. Atlantic	AMT24_09	RMNH.MOL.340313	SAMN21503571	OK185230	7507590	3882192	51.7	93.3	264.26
Lbul_AMT24_09_15	N. Atlantic	AMT24_09	RMNH.MOL.340314	SAMN21503572	OK185231	11013347	5403990	49.1	91.1	356.17
Lbul_AMT24_11_01	Eq. Atlantic	AMT24_11	RMNH.MOL.445261	SAMN21503573	OK185232	15025619	3134372	20.9	98	229.41
Lbul_AMT24_11_02	Eq. Atlantic	AMT24_11	RMNH.MOL.445262	SAMN21503574	OK185233	10155361	2082906	20.5	98.3	153.3
Lbul_AMT24_11_03	Eq. Atlantic	AMT24_11	RMNH.MOL.445263	SAMN21503575	OK185234	9435100	3723626	39.5	96.3	267.53
Lbul_AMT24_11_04	Eq. Atlantic	AMT24_11	RMNH.MOL.445264	SAMN21503576	OK185235	11671099	4537034	38.9	96.8	327.32
Lbul_AMT24_11_05	Eq. Atlantic	AMT24_11	RMNH.MOL.445265	SAMN21503577	OK185236	7025075	2885862	41.1	96	205.93
Lbul_AMT24_11_06	Eq. Atlantic	AMT24_11	RMNH.MOL.445266	SAMN21503578	OK185237	16388064	5543770	33.8	96.7	398
Lbul_AMT24_11_07	Eq. Atlantic	AMT24_11	RMNH.MOL.445267	SAMN21503579	OK185238	11967424	2397416	20.0	98.2	175.51
Lbul_AMT24_11_08	Eq. Atlantic	AMT24_11	RMNH.MOL.445268	SAMN21503580	OK185239	8415968	3502738	41.6	96.6	252.45
Lbul_AMT24_11_09	Eq. Atlantic	AMT24_11	RMNH.MOL.445269	SAMN21503581	OK185240	10391015	4210460	40.5	97	304.95
Lbul_AMT24_15_01	Eq. Atlantic	AMT24_15	RMNH.MOL.445270	SAMN21503582	OK185241	13380637	2561854	19.1	98.3	187.81
Lbul_AMT24_15_02	Eq. Atlantic	AMT24_15	RMNH.MOL.445271	SAMN21503583	OK185242	10118026	3776912	37.3	96.7	271.57
Lbul_AMT24_15_03	Eq. Atlantic	AMT24_15	RMNH.MOL.445272	SAMN21503584	OK185243	9231365	3472194	37.6	95.8	248.58
Lbul_AMT24_15_04	Eq. Atlantic	AMT24_15	RMNH.MOL.445273	SAMN21503585	OK185244	8614057	3267500	37.9	95.8	234.01
Lbul_AMT24_15_05	Eq. Atlantic	AMT24_15	RMNH.MOL.445274	SAMN21503586	OK185245	8749292	3394010	38.8	96.5	244.8
Lbul_AMT24_15_06	Eq. Atlantic	AMT24_15	RMNH.MOL.445275	SAMN21503587	OK185246	11202968	4045168	36.1	95.3	287.66
Lbul_AMT24_15_07	Eq. Atlantic	AMT24_15	RMNH.MOL.445276	SAMN21503588	OK185247	11035258	3769568	34.2	95.9	268.6

TABLE S1 Continued.

Specimen ID	Population	Station	Museum voucher code	Biosample	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth
Lbul_AMT24_15_08	Eq. Atlantic	AMT24_15	RMNH.MOL.445277	SAMN21503589	OK185248	8327631	3349770	40.2	96.6	242.2
Lbul_AMT24_15_10	Eq. Atlantic	AMT24_15	RMNH.MOL.445278	SAMN21503590	OK185249	8418989	3086620	36.7	96.1	221.15
Lbul_AMT24_17_03	Eq. Atlantic	AMT24_17	RMNH.MOL.341275	SAMN20293115	MZ542566	8548346	2641394	30.9	86.5	162.13
Lbul_AMT24_17_05	Eq. Atlantic	AMT24_17	RMNH.MOL.340317	SAMN20293116	MZ542567	8619139	4035590	46.8	91.7	267.48
Lbul_AMT24_17_06	Eq. Atlantic	AMT24_17	RMNH.MOL.340318	SAMN20293117	MZ542568	9525743	5112712	53.7	93.6	348.92
Lbul_AMT24_17_07	Eq. Atlantic	AMT24_17	RMNH.MOL.340319	SAMN20293118	MZ542569	5945465	3128096	52.6	93	211.7
Lbul_AMT24_17_08	Eq. Atlantic	AMT24_17	RMNH.MOL.340320	SAMN20293119	MZ542570	13129266	7287498	55.5	88.8	467.92
Lbul_AMT24_17_09	Eq. Atlantic	AMT24_17	RMNH.MOL.340321	SAMN20293120	MZ542571	10888222	5872132	53.9	87.9	371.76
Lbul_AMT24_17_10	Eq. Atlantic	AMT24_17	RMNH.MOL.340322	SAMN20293121	MZ542572	13675828	7318800	53.5	87.2	458.84
Lbul_AMT24_17_12	Eq. Atlantic	AMT24_17	RMNH.MOL.340323	SAMN20293122	MZ542573	4963728	2581066	52.0	94	177.65
Lbul_AMT24_17_13	Eq. Atlantic	AMT24_17	RMNH.MOL.340324	SAMN20293123	MZ542574	10474291	5016032	47.9	91.3	331.99
Lbul_AMT24_17_14	Eq. Atlantic	AMT24_17	RMNH.MOL.340325	SAMN20293124	MZ542575	9865549	5068552	51.4	92	338.06
Lbul_AMT24_17_15	Eq. Atlantic	AMT24_17	RMNH.MOL.340326	SAMN20293125	MZ542576	8087831	3514198	43.5	91.3	231.62
Lbul_AMT24_18_05	Eq. Atlantic	AMT24_18	RMNH.MOL.340328	SAMN21503591	OK185250	8007571	3090614	38.6	89.8	198.8
Lbul_AMT24_18_06	Eq. Atlantic	AMT24_18	RMNH.MOL.340329	SAMN21503592	OK185251	9055257	3931836	43.4	91.1	258.63
Lbul_AMT24_18_07	Eq. Atlantic	AMT24_18	RMNH.MOL.340330	SAMN21503593	OK185252	8158759	4070016	49.9	93.8	277.72
Lbul_AMT24_18_08	Eq. Atlantic	AMT24_18	RMNH.MOL.340331	SAMN21503594	OK185253	7833978	3731652	47.6	93	253.14
Lbul_AMT24_18_09	Eq. Atlantic	AMT24_18	RMNH.MOL.340332	SAMN21503595	OK185254	7447577	4267036	57.3	89.1	275.02
Lbul_AMT24_18_10	Eq. Atlantic	AMT24_18	RMNH.MOL.340333	SAMN21503596	OK185255	1982112	1188186	59.9	89.7	77.18
Lbul_AMT24_18_11	Eq. Atlantic	AMT24_18	RMNH.MOL.340334	SAMN21503597	OK185256	9178026	4424250	48.2	85.9	273.14
Lbul_AMT24_18_13	Eq. Atlantic	AMT24_18	RMNH.MOL.340335	SAMN21503598	OK185257	8481962	4646702	54.8	93.9	319.62
Lbul_AMT24_18_14	Eq. Atlantic	AMT24_18	RMNH.MOL.340336	SAMN21503599	OK185258	11335305	5722314	50.5	91.9	381.68
Lbul_AMT24_18_15	Eq. Atlantic	AMT24_18	RMNH.MOL.340337	SAMN21503600	OK185259	9547302	4829890	50.6	92.6	325.75
Lbul_AMT24_20_03	S. Atlantic	AMT24_20	RMNH.MOL.341197	SAMN11131474	OK185260	9080282	4662318	51.3	92.4	314.2
Lbul_AMT24_20_04	S. Atlantic	AMT24_20	RMNH.MOL.341198	SAMN11131475	OK185261	4924010	1294918	26.3	86.8	79.7
Lbul_AMT24_20_06	S. Atlantic	AMT24_20	RMNH.MOL.340339	SAMN21503601	OK185262	17306981	7992980	46.2	89.9	517.1
Lbul_AMT24_20_07	S. Atlantic	AMT24_20	RMNH.MOL.340340	SAMN21503602	OK185263	7406817	2811576	38.0	92.1	186.88
Lbul_AMT24_20_08	S. Atlantic	AMT24_20	RMNH.MOL.340341	SAMN21503603	OK185264	10531079	5480172	52.0	87.5	344.98
Lbul_AMT24_20_09	S. Atlantic	AMT24_20	RMNH.MOL.341199	SAMN11131476	OK185265	9231573	4246804	46.0	85.9	261.19
Lbul_AMT24_20_10	S. Atlantic	AMT24_20	RMNH.MOL.340342	SAMN21503604	OK185266	21061787	9944826	47.2	87.8	630.7
Lbul_AMT24_20_11	S. Atlantic	AMT24_20	RMNH.MOL.340343	SAMN21503605	OK185267	11309440	5591156	49.4	88.5	357.81
Lbul_AMT24_20_12	S. Atlantic	AMT24_20	RMNH.MOL.340344	SAMN21503606	OK185268	9597819	3747444	39.0	88.9	239.05
Lbul_AMT24_22_04	S. Atlantic	AMT24_22	RMNH.MOL.341200	SAMN11131477	MZ542577	16895753	4670866	27.6	85.6	284.34
Lbul_AMT24_22_06	S. Atlantic	AMT24_22	RMNH.MOL.341201	SAMN11131478	MZ542578	6107748	1909732	31.3	82.6	111.5
Lbul_AMT24_22_07	S. Atlantic	AMT24_22	RMNH.MOL.340345	SAMN20293126	MZ542579	12151985	6428490	52.9	92.4	432.12
Lbul_AMT24_22_08	S. Atlantic	AMT24_22	RMNH.MOL.341202	SAMN11131479	MZ542580	8193775	3492314	42.6	92.5	234.08
Lbul_AMT24_22_09	S. Atlantic	AMT24_22	RMNH.MOL.340346	SAMN20293127	MZ542581	7545267	4085630	54.1	94	280.28
Lbul_AMT24_22_11	S. Atlantic	AMT24_22	RMNH.MOL.340348	SAMN20293128	MZ542582	12123113	6301860	52.0	87.3	398.85

TABLE S1 Continued.

Specimen ID	Population	Station	Museum voucher code	Biosample	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth
Lbul_AMT24_22_12	S. Atlantic	AMT24_22	RMNH.MOL.340349	SAMN20293129	MZ542583	16218485	8033820	49.5	87.9	512.47
Lbul_AMT24_22_13	S. Atlantic	AMT24_22	RMNH.MOL.340350	SAMN20293130	MZ542584	12379217	6530426	52.8	89.1	422.73
Lbul_AMT24_22_14	S. Atlantic	AMT24_22	RMNH.MOL.340351	SAMN20293131	MZ542585	9636983	4995398	51.8	92.4	336.15
Lbul_AMT24_22_15	S. Atlantic	AMT24_22	RMNH.MOL.340352	SAMN20293132	MZ542586	10732523	5037788	46.9	91.2	332.12
Lbul_AMT24_23_02	S. Atlantic	AMT24_23	RMNH.MOL.340353	SAMN20293133	MZ542587	10912937	2885556	26.4	87.7	179.76
Lbul_AMT24_23_05	S. Atlantic	AMT24_23	RMNH.MOL.340354	SAMN20293134	MZ542588	9060457	4416126	48.7	91.7	292.49
Lbul_AMT24_23_06	S. Atlantic	AMT24_23	RMNH.MOL.341203	SAMN11131480	MZ542589	10248846	3785256	36.9	83.4	223.16
Lbul_AMT24_23_07	S. Atlantic	AMT24_23	RMNH.MOL.340355	SAMN20293135	MZ542590	12388848	6365196	51.4	92.6	428.94
Lbul_AMT24_23_08	S. Atlantic	AMT24_23	RMNH.MOL.341204	SAMN11131481	MZ542591	13640477	6294478	46.1	86.4	390.4
Lbul_AMT24_23_09	S. Atlantic	AMT24_23	RMNH.MOL.341205	SAMN11131482	MZ542592	14038163	6982544	49.7	86.7	434.67
Lbul_AMT24_23_10	S. Atlantic	AMT24_23	RMNH.MOL.340356	SAMN20293136	MZ542593	10848376	6083488	56.1	88.8	390.94
Lbul_AMT24_23_11	S. Atlantic	AMT24_23	RMNH.MOL.340357	SAMN20293137	MZ542594	8918248	4643746	52.1	93.7	317.31
Lbul_AMT24_23_12	S. Atlantic	AMT24_23	RMNH.MOL.340358	SAMN20293138	MZ542595	15044933	6750610	44.9	90.3	440.47
Lbul_AMT24_23_13	S. Atlantic	AMT24_23	RMNH.MOL.340359	SAMN20293139	MZ542596	12080446	5990002	49.6	91.3	396.18
Lbul_NIC2_S1C3_01	N. Atlantic	NIC2_S1	RMNH.MOL.346557	SAMN20293216	MZ542673	17269646	5816130	33.7	96.5	418.65
Lbul_NIC2_S1C3_02	N. Atlantic	NIC2_S1	RMNH.MOL.346558	SAMN20293217	MZ542674	11362364	4261032	37.5	96.2	305.98
Lbul_NIC2_S1C3_03	N. Atlantic	NIC2_S1	RMNH.MOL.346559	SAMN20293218	MZ542675	13073572	2907762	22.2	98.2	214.8
Lbul_NIC2_S1C3_07	N. Atlantic	NIC2_S1	RMNH.MOL.346560	SAMN20293219	MZ542676	19015434	3526366	18.5	98.3	258.02
Lbul_NIC2_S1C3_08	N. Atlantic	NIC2_S1	RMNH.MOL.346561	SAMN20293220	MZ542677	12518723	4716858	37.7	95	333.98
Lbul_NIC2_S1C3_10	N. Atlantic	NIC2_S1	RMNH.MOL.346562	SAMN20293221	MZ542678	11557090	4428220	38.3	96.1	317.52
Lbul_NIC2_S1C3_11	N. Atlantic	NIC2_S1	RMNH.MOL.346563	SAMN20293222	MZ542679	9691816	3957806	40.8	97.1	288.39
Lbul_NIC2_S1C3_12	N. Atlantic	NIC2_S1	RMNH.MOL.346564	SAMN20293223	MZ542680	12914232	5011224	38.8	96.2	359.34
Lbul_NIC2_S1C3_13	N. Atlantic	NIC2_S1	RMNH.MOL.346565	SAMN20293224	MZ542681	12934158	2699656	20.9	97.8	197.4
Lbul_NIC2_S1C3_14	N. Atlantic	NIC2_S1	RMNH.MOL.346566	SAMN20293225	MZ542682	6773313	2628106	38.8	96.8	190
Lbul_NIC2_S1C3_15	N. Atlantic	NIC2_S1	RMNH.MOL.346567	SAMN20293226	MZ542683	15422685	4556206	29.5	94.2	317.07
Lbul_NIC2_S9C3_01	Eq. Atlantic	NIC2_S9	RMNH.MOL.346568	SAMN20293227	MZ542684	9518473	3873588	40.7	96.9	280.88
Lbul_NIC2_S9C3_02	Eq. Atlantic	NIC2_S9	RMNH.MOL.346569	SAMN20293228	MZ542685	12754807	2595230	20.3	98.4	190.72
Lbul_NIC2_S9C3_03	Eq. Atlantic	NIC2_S9	RMNH.MOL.346570	SAMN20293229	MZ542686	12515237	2553328	20.4	98.2	186.96
Lbul_NIC2_S9C3_04	Eq. Atlantic	NIC2_S9	RMNH.MOL.346571	SAMN20293230	MZ542687	11723092	2511120	21.4	97.9	183.6
Lbul_NIC2_S9C3_05	Eq. Atlantic	NIC2_S9	RMNH.MOL.346572	SAMN20293231	MZ542688	6621308	2916484	44.0	96.1	209.55
Lbul_NIC2_S9C3_06	Eq. Atlantic	NIC2_S9	RMNH.MOL.346573	SAMN20293232	MZ542689	16042144	3317304	20.7	98.2	243.47
Lbul_NIC2_S9C3_07	Eq. Atlantic	NIC2_S9	RMNH.MOL.346574	SAMN20293233	MZ542690	11817232	5046408	42.7	95.4	358.51
Lbul_NIC2_S9C3_08	Eq. Atlantic	NIC2_S9	RMNH.MOL.346575	SAMN20293234	MZ542691	8288640	3383882	40.8	96.8	245.5
Lbul_NIC2_S9C3_09	Eq. Atlantic	NIC2_S9	RMNH.MOL.346576	SAMN20293235	MZ542692	9531410	4113674	43.2	96.4	296.07
Lbul_NIC2_S9C3_10	Eq. Atlantic	NIC2_S9	RMNH.MOL.346577	SAMN20293236	MZ542693	8715280	3681980	42.2	97	266.83
Lbul_NIC8_S2C3-4_04	N. Atlantic	NIC8_S2	RMNH.MOL.445279	SAMN21503607	OK185269	8420748	3564880	42.3	97.2	258.92
Lbul_NIC8_S2C3-4_05	N. Atlantic	NIC8_S2	RMNH.MOL.445280	SAMN21503608	OK185270	13328600	2839560	21.3	98.1	208.45
Lbul_NIC8_S2C3-4_06	N. Atlantic	NIC8_S2	RMNH.MOL.445281	SAMN21503609	OK185271	14532267	5963484	41.0	95.7	425.32

TABLE S1 Continued.

Specimen ID	Population	Station	Museum voucher code	Biosample	COI	Raw reads	Final mapped reads	% HQ mapped	% HQ reads on target	Depth
Lbul_NIC8_S2C3-4_08	N. Atlantic	NIC8_S2	RMNH.MOL.445282	SAMN21503610	OK185272	10461231	4573766	43.7	95.9	326.46
Lbul_NIC8_S2C3-4_09	N. Atlantic	NIC8_S2	RMNH.MOL.445283	SAMN21503611	OK185273	8520547	3611620	42.4	96.5	261.14
Lbul_NIC8_S2C3-4_10	N. Atlantic	NIC8_S2	RMNH.MOL.445284	SAMN21503612	OK185274	11279836	2381762	21.1	98	174.82
Lbul_NIC8_S2C3-4_11	N. Atlantic	NIC8_S2	RMNH.MOL.445285	SAMN21503613	OK185275	8042805	3441228	42.8	96.8	249.87
Lbul_NIC8_S2C3-4_12	N. Atlantic	NIC8_S2	RMNH.MOL.445286	SAMN21503614	OK185276	12862527	4794738	37.3	95	338.95
Lbul_NIC8_S2C3-4_13	N. Atlantic	NIC8_S2	RMNH.MOL.445287	SAMN21503615	OK185277	8389718	3390144	40.4	96.2	243.71
Lbul_NIC8_S2C3-4_15	N. Atlantic	NIC8_S2	RMNH.MOL.445288	SAMN21503616	OK185278	9223142	3876222	42.0	96.8	281.04
Lbul_NIC8_S4C1-2_01	N. Atlantic	NIC8_S4	RMNH.MOL.445289	SAMN21503617	OK185279	11970090	4312268	36.0	96	308.62
Lbul_NIC8_S4C1-2_02	N. Atlantic	NIC8_S4	RMNH.MOL.445290	SAMN21503618	OK185280	13872652	5280140	38.1	95.5	377.21
Lbul_NIC8_S4C1-2_03	N. Atlantic	NIC8_S4	RMNH.MOL.445291	SAMN21503619	OK185281	9435957	3658778	38.8	96	262.38
Lbul_NIC8_S4C1-2_04	N. Atlantic	NIC8_S4	RMNH.MOL.445292	SAMN21503620	OK185282	11716495	4275412	36.5	96.2	307.27
Lbul_NIC8_S4C1-2_05	N. Atlantic	NIC8_S4	RMNH.MOL.445293	SAMN21503621	OK185283	14562144	5354760	36.8	94.6	377.49
Lbul_NIC8_S4C1-2_06	N. Atlantic	NIC8_S4	RMNH.MOL.445294	SAMN21503622	OK185284	15928479	2872534	18.0	98.7	211.53
Lbul_NIC8_S4C1-2_07	N. Atlantic	NIC8_S4	RMNH.MOL.445295	SAMN21503623	OK185285	15264834	6241996	40.9	96.2	447.08
Lbul_NIC8_S4C1-2_08	N. Atlantic	NIC8_S4	RMNH.MOL.445296	SAMN21503624	OK185286	10253878	3644592	35.5	94.4	255.05
Lbul_NIC8_S4C1-2_09	N. Atlantic	NIC8_S4	RMNH.MOL.445297	SAMN21503625	OK185287	8930184	3495348	39.1	96.2	251.3
Lbul_NIC8_S4C1-2_10	N. Atlantic	NIC8_S4	RMNH.MOL.445298	SAMN21503626	OK185288	14551896	2907852	20.0	97.7	212.1
Lbul_NIC8_S4C1-2_12	N. Atlantic	NIC8_S4	RMNH.MOL.445299	SAMN21503627	OK185289	10861589	4261284	39.2	96.4	307.42
Lbul_NIC8_S5C3_01	N. Atlantic	NIC8_S5	RMNH.MOL.346578	SAMN20293237	MZ542694	7064864	2931840	41.5	96.4	211.49
Lbul_NIC8_S5C3_03	N. Atlantic	NIC8_S5	RMNH.MOL.346579	SAMN20293238	MZ542695	8865106	3757808	42.4	96.4	271.51
Lbul_NIC8_S5C3_04	N. Atlantic	NIC8_S5	RMNH.MOL.346580	SAMN20293239	MZ542696	14846520	5444886	36.7	94.6	383.93
Lbul_NIC8_S5C3_06	N. Atlantic	NIC8_S5	RMNH.MOL.346581	SAMN20293240	MZ542697	12715480	5027460	39.5	96.4	362.09
Lbul_NIC8_S5C3_07	N. Atlantic	NIC8_S5	RMNH.MOL.346582	SAMN20293241	MZ542698	11893813	4701916	39.5	96.6	339.21
Lbul_NIC8_S5C3_08	N. Atlantic	NIC8_S5	RMNH.MOL.346583	SAMN20293242	MZ542699	16832853	6087524	36.2	95.6	434.87
Lbul_NIC8_S5C3_09	N. Atlantic	NIC8_S5	RMNH.MOL.346584	SAMN20293243	MZ542700	10162231	2142102	21.1	97.2	155.04
Lbul_NIC8_S5C3_11	N. Atlantic	NIC8_S5	RMNH.MOL.346585	SAMN20293244	MZ542701	8608429	2917080	33.9	95.4	206.51
Lbul_NIC8_S5C3_12	N. Atlantic	NIC8_S5	RMNH.MOL.346586	SAMN20293245	MZ542702	14259468	5927612	41.6	96.2	424.53
Lbul_NIC8_S5C3_13	N. Atlantic	NIC8_S5	RMNH.MOL.346587	SAMN20293246	MZ542703	6800160	2862764	42.1	97.1	208.5
Lbul_NIC8_S5C3_14	N. Atlantic	NIC8_S5	RMNH.MOL.346588	SAMN20293247	MZ542704	8123312	3435518	42.3	96.5	248.52

6

Creativity in times of COVID-19: Using hand-sanitiser for morphometric and genetic analysis of zooplankton

L.Q. Choo, G. Spagliardi, K.T.C.A. Peijnenburg

Abstract

There is a lack of standardized photography methods for marine zooplankton and other small organisms due to the difficulty in manipulating such small and often fragile specimens. Yet, standardized photographs provide important morphological information to accompany DNA-barcoded specimens, and allow for in-depth morphometric analyses. We used alcohol-based hand sanitiser, which became widely available during the COVID-19 crisis, as a medium for photographing pteropods of the genus *Limacina* prior to micro-CT scanning and destructive DNA analysis. The high viscosity and transparency of the hand sanitiser enabled easy handling of the specimens so that they could be positioned in a standardised orientation and photographed with a stacking microscope. The high quality photographs provide a record of morphology and allow for subsequent geometric morphometric analyses. This method did not impact the downstream micro-CT and molecular analyses and resulted in ten reference DNA barcodes. When alcohol-based hand sanitiser entered our daily lives due to the COVID-19 pandemic, we learned that we could make use of it as a cheap and easily available resource to make high quality voucher photographs of zooplankton. This serves as a record of the morphology of the organism for taxonomic identification and facilitates further studies involving morphology, growth and ontogeny.

Need for high quality voucher photographs

The COVID-19 crisis has disrupted laboratory work globally with severe restrictions to working in the lab, as well as shortages in reagents and disposables. In spite of that, it has also provided opportunities for change and innovation. Alcohol-based hand sanitiser became increasingly more common since the start of the pandemic as a convenient replacement for soap and water to maintain hand hygiene. Because of the availability of alcohol-based hand sanitiser, we were inspired to include it as part of our integrative taxonomy workflow to facilitate the photography of small pteropods.

Shelled pteropods are holoplanktonic gastropods that are a common subject of global change research, due to their important ecological and biogeochemical roles and susceptibility to ocean acidification. Shelled pteropods have been regarded as bio-indicators of ocean acidification, because their thin aragonitic shell is sensitive to dissolution under acidified conditions (Bednaršek et al., 2017b; Orr et al., 2005). Pteropods contribute to the oceanic carbon flux by producing biomass and sequestering carbon through high phytoplankton grazing (Hunt et al., 2008) and large downward fluxes of faecal pellets (Manno et al., 2010), mucus nets (Conley et al., 2018; Noji et al., 1997), and shells (Fabry et al., 2009; Tsurumi et al., 2005). Pteropods are also prey for heteropods (Böer et al., 2005), amphipods (Bernard, 2006), cephalopods (Hanlon and Messenger, 1998), and, in polar systems, fishes, seabirds (Hunt et al., 2008) and marine mammals (Lalli and Gilmer, 1989). Their species diversity and population structure is commonly assessed with DNA barcodes (e.g., Burridge et al., 2017b; Choo et al., 2021; Hunt et al., 2010; Jennings et al., 2010; Kohnert et al., 2020), which contributes to a global zooplankton DNA barcoding reference database.

There is an urgent need for more high quality voucher photographs of specimens to complement DNA barcode reference databases of marine zooplankton (Bucklin et al., 2021b). This allows for the morphological identification of specimens, and validation of their DNA barcodes (Bucklin et al., 2016; Laakmann et al., 2020). The preservation of morphological characters of the specimen, in addition to sequencing of DNA barcodes, and collection of associated georeferencing, environmental and ecological metadata, are important components of an integrative taxonomy pipeline (Padial et al., 2010). With the aid of these multiple sources of information, we can successfully assess species boundaries within the zooplankton (Bode et al., 2017; Burridge et al., 2019; Hirai et al., 2015), and in some cases, resolve (pseudo-)cryptic species identified through DNA barcodes with more thorough inspection of their morphology (e.g., Wall-Palmer et al., 2018). While non-destructive DNA extraction is possible in some zooplankton taxa by subsampling (e.g., Choquet et al., 2018), or by leaching DNA without damage to calcite shells or chitinous exoskeletons (e.g., Cornils, 2015; Weiner et al., 2016), in most cases, the small size of zooplankton means that they are destroyed during DNA extraction protocols. In these cases, high quality voucher photographs play a critical role as the only remaining record of the morphological traits of the barcoded specimen.

There is a lack of methodological detail as to how zooplankton can be prepared for taking a high quality stacking photograph. Ideally, these specimens should be positioned in a standardised orientation to facilitate measurements of body parts, or to highlight a particular aspect of their anatomy. In coiled gastropods especially, the positioning of the specimen is important to standardise the turn and tilt of the shell (Callomon, 2019), which directly impacts the suitability of the photographs in geometric morphometrics analyses. Known methods for positioning small invertebrates for photography vary from malleable tack for dry specimens to glass slides, wax cradles and stainless steel nuts for wet specimens (Geiger et al., 2007). Micro-computed tomography (micro-CT) scanning is a method where specimens do not need to be placed in a standardised orientation because 3D reconstructions of their shell morphology are obtained (Shimizu et al., 2018, 2021), but it is expensive and time-consuming to have 3D scans for all specimens that are included in the DNA barcoding pipeline.

As part of our integrative taxonomy pipeline for planktonic gastropods, we needed to find a reproducible way to image specimens in a standardised orientation before destructive DNA extraction. Alcohol-based hand sanitiser has been commonly used in positioning and stabilising spiders (e.g., Bilton, 2018; LeMay and Agnarsson, 2020; Valdez-Mondragón, 2010) and other macroinvertebrates (http://www.gigamacro.com/blog/wet-specimen-holder-macroinvertebrates) for photography, as well as in suspending preserved macroinvertebrates in glass vials for display and outreach activities (https://thedragonflywoman.com/2011/02/21/hand-sanitizer-preservation/). The jelly-like texture of hand sanitiser allows for manipulation of the specimen while simultaneously providing enough viscosity for the specimen to remain in the desired position. To the best of our knowledge, this is the first documentation of alcohol-based hand sanitiser in zooplankton photography as part of a DNA barcoding pipeline.

Use of hand-sanitiser to orient and photograph small organisms

We used alcohol-based hand sanitiser in the photography of ten shelled pteropods belonging to the *Limacina* genus, representing five nominal species: *L. helicina*, *L. retroversa*, *L. bulimoides*, *L. trochiformis*, *L. lesueuerii*. Alcohol-based hand sanitiser was used as a positioning medium for photographing the specimen in a standardised apertural orientation. The store-bought hand sanitiser we used ('Dr. Original' brand) contained the following ingredients: Denatured alcohol, Water, Polyacrylamide, C13-14 Isoparaffin and Laureth-7.

Under the microscope, we prepared the setup for imaging, which consisted of a watch glass with a small volume of hand sanitiser, atop on a black velvet cloth which provided a dark background (Figure 1A). Each specimen was gently placed on the blob of hand sanitiser in a watch glass. The specimen was then manipulated carefully into a standard orientation with a thin brush or stick, to ensure the shell

axis was parallel to the plane of the camera (Figure 1B). The specimen and hand sanitiser were then covered with a thin layer of 96% ethanol to reduce glare (Figure 1C). Subsequently, the shells were photographed using a Zeiss V20 stacking stereomicroscope with Axiovision software (Zeiss, Germany). Once photographed, each specimen was rinsed in 96% ethanol and stored in 96% ethanol at -20 °C.

After the photograph was obtained for each specimen, a subset of five individuals were used for micro-CT scanning. Specimens were prepared and scanned according to the protocol described in (Mekkes et al., 2021a). We also obtained the mitochondrial Cytochrome Oxidase I (COI) barcode of the 10 individuals. DNA was extracted and barcoded according to the protocol in Choo et al. (2021). The shells were washed in 96% ethanol before micro-CT scanning to prevent possible artefacts from the gel, and in water before being placed in lysis buffer for DNA extraction to prevent possible inhibition of the polymerase chain reaction (PCR) due to any remaining (denatured) alcohol. Sequences were checked and edited in Geneious Prime 2021.1.1 (https://www.geneious.com) to produce the final sequences that were submitted to the BOLD database (Ratnasingham and Hebert, 2007, 2013).

Integrative taxonomy of shelled pteropods

Alcohol-based hand sanitiser facilitates the imaging process of ethanol-preserved shelled pteropods and does not interfere with their downstream micro-CT scanning, DNA extraction, and barcoding. We experienced that orienting the specimens for stacking photography was efficient with alcohol gel and the transparent shells contrasted well against the clean dark background provided by the black velvet cloth. These high quality images can be used for publication and geometric morphometric analyses and were included as digital vouchers in a reference database with minimal editing needed (Figure 2).

The DNA extraction and COI barcoding of the 10 specimens were successful, including those that had the additional step of micro-CT scanning prior to DNA

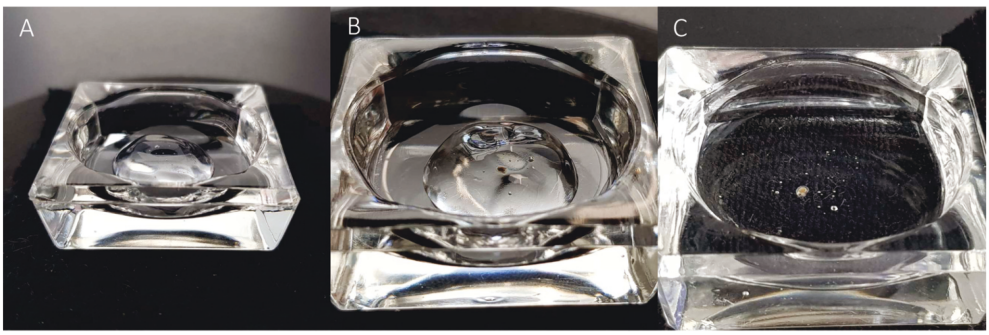


Figure 1 Imaging setup for *Limacina* specimens. (A) A glass watch glass was placed on a sheet on black velvet under the microscope, and hand sanitiser was added within the watch glass. (B) The specimen was placed within the hand sanitiser and manipulated into position. (C) A small volume of 96% ethanol was added to completely cover the hand sanitiser.

extraction. This adds further evidence to initial exploratory experiments (Hall et al., 2014) that DNA barcoding is compatible with micro-CT scanning and barcodes can be reliably amplified despite the radiation from the micro-CT. The DNA vouchers of these specimens are stored and accessible through Naturalis Biodiversity Center (RMNH.MOL347126, 347129, 347279-280, 347361, 347363, 347367-68, 347375, 347382), and their 2D photographs, 3D micro-CT reconstructions (where applicable) and COI barcodes were submitted to BOLD (BOLD accessions: LPCOI001-21-LPCOI010-21).

Towards better reference databases for marine zooplankton

This method of pteropod photography with alcohol-based hand sanitiser has the potential to be more broadly applied to the barcoding pipelines of other marine zooplankton taxa. We expect that the increased flexibility in positioning specimens with hand sanitiser can allow for more accurate measurements of appendages in planktonic animals such as copepods, euphausiids and other crustaceans, as well as facilitate the orientation of specimens for observation of diagnostic traits or the re-assembly of fragile species that fall apart upon collection, e.g., siphonophores for photography. The gelatinous medium also prevents any friction of the specimens against a solid surface, substantially decreasing the risk of damaging the fragile organisms while orienting them.

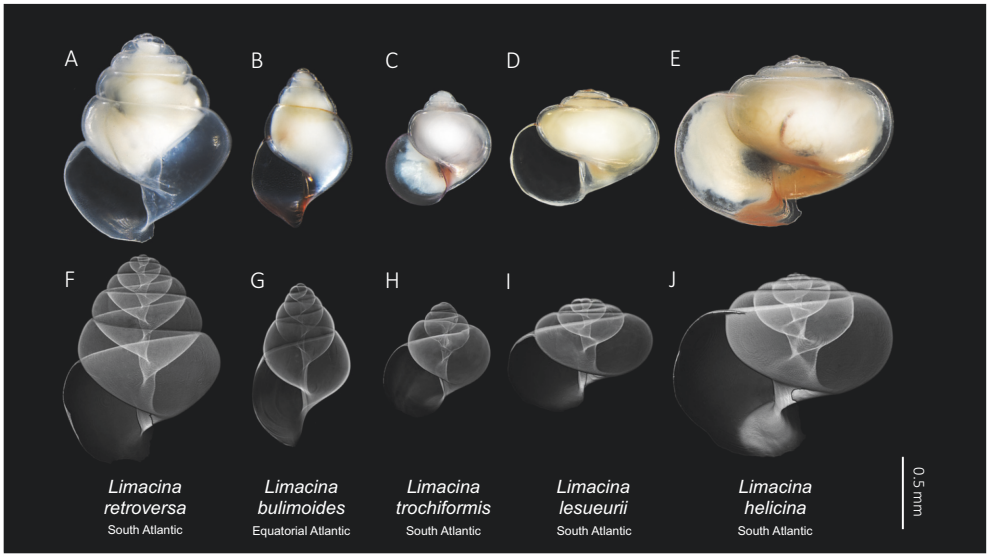


Figure 2 Stacking photographs (top row) and micro-CT 3D reconstructions (bottom row) of one specimen from each of the five nominal *Limacina* species. (A, F) *L. retroversa*, (B, G) *L. bulimoides*, (C, H) *L. trochiformis*, (D, I) *L. lesueurii*, (E, J) *L. helicina*. The stacking photographs and micro-CT reconstructions for each species are of the same individual. The species name and collection locality are stated beneath each individual.

We have only tested one brand of hand sanitiser in our protocol. Hand sanitiser typically contains the active ingredients ethanol, isopropyl alcohol or n-isopropanol, and inert ingredients like water, humectants, and stabilisers (Golin et al., 2020), although the exact compositions may vary by brand. The different compositions of ingredients may affect their viscosity and transparency to a large degree (Berardi et al., 2020). Both viscosity and transparency are important traits that impact the stability of the specimen after manipulating it into position, and the image quality of the photograph. We note that the viscosity of the gel can be increased by decreasing its temperature (e.g., by keeping the hand sanitiser in the fridge). Trial and error for testing the suitability of different store-bought hand sanitiser brands is necessary. Alcohol-based hand sanitisers also vary in terms of their alcohol concentrations, between the range of 60-95% (Boyce and Pittet, 2002; Edmonds et al., 2012). We have not tested the effect of differing alcohol concentrations in hand sanitiser, but expect that the exact ethanol concentration of the alcohol gel matters less than the initial fixation quality of the specimen when it comes to subsequent DNA extraction and PCR (Stein et al., 2013).

We hope the described methodology can increase the efficiency and quality of 2D stacking photography and encourage the inclusion of digital vouchers in DNA barcode reference databases of marine zooplankton. It is important to have such voucher photographs of specimens, especially if specimens have to be destroyed for DNA extraction (Bucklin et al., 2021b). In this way, although the specimen is no longer available after barcoding, it is still possible to link genetic information with phenotypic characters through high quality photographs and 3D reconstructions. The ease of positioning specimens in a standard orientation also facilitates the challenging photography of younger life stages, such as veliger larvae in gastropods and nauplii in copepods, prior to DNA extraction. This is necessary to provide a complete overview of morphology across the life history of the species, facilitate species identifications and resolve cryptic species complexes (Bode et al., 2017). While alcohol-based hand sanitiser entered our daily lives due to a distressing pandemic, we learned that we can make use of it as a cheap and easily available resource as part of an integrative DNA barcoding pipeline of marine zooplankton.

Acknowledgements

This study was supported by a Netherlands Organisation for Scientific Research (NWO) Vidi grant 016.161.351 to K.T.C.A.P. We wish to warmly thank Deborah Wall-Palmer, Anders Illum and Nikolaj Scharff for their insights leading to our use of alcohol-based hand sanitiser in pteropod photography, as well as Luis Martell for discussion about its use for hydrozoan imaging. We are also grateful to Jef Huisman and Galice Hoarau for comments on the manuscript and Frank Stokvis and Rob Langelaan at Naturalis Biodiversity Center for their help and expertise in implementing the DNA barcoding and micro-CT scanning pipelines.

Data availability

The data underlying this article are available in BOLD at http://www.boldsystems. org/, and can be accessed with BOLD Process IDs LPCOI001-21 to LPCOI010-21. The DNA vouchers

(RMNH.MOL347126, 347129, 347279-280, 347361, 347363, 347367-68, 347375, 347382) are accessioned at Naturalis Biodiversity Center, the Netherlands, and available upon request.

7

General discussion

To gain insight into the population structure, biogeographical distribution and evolutionary potential of pteropods, I studied the circumglobal pteropod *Limacina bulimoides* using barcoding genes and genome-wide markers. The relative lack of hard physical barriers in the marine realm has led to previous expectations of genetically homogeneous populations across the wide geographical ranges of marine species, especially in holoplankton that are characterized by large effective population sizes and high dispersal potentials (Allendorf et al., 2010; Norris, 2000; Palumbi, 1994; Peijnenburg and Goetze, 2013; van der Spoel and Heyman, 1983). However, I found genetically distinct and geographically separated lineages located in the Atlantic, Indo-Pacific and Pacific Ocean (Chapter 4) and populations separated by dispersal barriers within the Atlantic basin (Chapters 3, 5). This is consistent with numerous other studies of marine species where populations are more structured than expected (e.g. Levin, 2006; Palumbi, 2004; Peijnenburg and Goetze, 2013), or where cryptic species hiding in plain sight are uncovered with the aid of genetic techniques (Bucklin et al., 2018; Halbert et al., 2013; Leray and Knowlton, 2016).

In my thesis, the observed genetic structure of *L. bulimoides* was assessed in relation to their ecology, as well as their morphological variation, which can be studied more easily through improvements in microscopic imaging techniques (Chapter 6). The results enabled characterisation of species boundaries and dispersal barriers in *L. bulimoides* and point to potential drivers of speciation in shelled pteropods. This contributes to our understanding of their ability to respond to future ocean changes.

In the following discussion, I summarise my findings from across the different chapters and point to future research directions, based on the following themes:

1. Added perspectives from genome-wide variability
2. Reliability of the COI barcode
3. Assessing barriers to dispersal using genetic and morphometric variation
4. Towards identifying the nature of dispersal barriers in the Atlantic Ocean
5. Species boundaries in the open ocean
6. Potential to adapt to future ocean conditions

Added perspectives from genome-wide variability

With the additional information obtained from genome-wide probes, a broader perspective of the evolutionary history of organisms is now possible as compared to earlier approaches based on the use of a limited number of barcoding genes. Genome-wide target capture probes can provide multiple independent observations of the evolutionary history of an organism, as diploid nuclear genomes are subject to crossing over and recombination during replication. With these multiple observations (1000s to 100 000s of single nucleotide polymorphisms, SNPs), the statistical strength of analyses increases, which facilitates the detection of subtle

population structure across organisms with high potential for gene flow (Waples, 1998). Hence, genome-wide analyses are more suitable for detecting population structure in oceanic circumglobal species, where genetic differentiation can occur over a range of spatial scales (Hellberg, 2009; Weersing and Toonen, 2009).

With genome-wide probes, we can also better determine whether hybridisation and gene flow have occurred between populations. In Chapter 2, discordant signals between the sequenced fragments of the mitochondrial COI gene and their sampling locations were recorded, which were identified as expatriates. However, with genome-wide data, I could ascertain that it was likely to be mitochondrial introgression as these individuals showed nuclear genomic variation that matched with their sampling location (Chapter 5). The introgression appears to be bidirectional across the South Atlantic subtropical gyre, as identified with more comprehensive sampling across individuals in Chapter 2. Introgression across the dispersal barrier in the South Atlantic subtropical gyre could indicate historical hybridisation of both populations adjacent to the barrier. Introgression is more commonly found to be unidirectional (Toews and Brelsford, 2012), and the occurrence of bidirectional introgression in *L. bulimoides* may indicate distinct selective pressures or environmental preferences. We would require more information on their ecology, behaviour and demography to identify the factors responsible for this pattern of introgression.

The multiple regions that are studied across the genome can also be used to avoid confounding effects when barcoding genes from mitochondrial and nuclear genomes (e.g., COI and 28S) have discordant histories and it is unclear which gives a more accurate representation of the species' evolutionary history (Chapter 2). Discordance between mitochondrial and nuclear genomes can indicate complex demographic and evolutionary scenarios, because the two genomes do not respond equally to demographic fluctuations (Després, 2019; Hinojosa et al., 2019). Hence, it is important to take information from both genomes into account. Genome-wide regions also provide more detail on the population structure in comparison to barcoding genes, because barcoding genes are often functionally constrained (Stoeckle and Thaler, 2014). One example is the 28S gene in *L. bulimoides* that did not distinguish between the three major genetic lineages (Chapter 2), which could be identified with confidence based on genome-wide SNPs (Chapter 4).

While the genome-wide probes provided more insight compared to barcoding genes, the success of the target capture method varied according to the different types of probes that were designed. On the whole, the most successful probes were those of the nuclear coding regions, while probes of the non-coding and mitochondrial regions did not fare so well. The sole mitochondrial region that was recovered consistently across individuals with high coverage was the COI barcode region designed from a Sanger sequenced region instead of a putative mitochondrial contig from the draft genome (Chapter 3). Mixed probe designs combining

both nuclear and organelle regions may encounter problems in recovering mitochondrial DNA reliably (Choquet, 2017). Nonetheless, it is also possible that the putative mitochondrial contig identified from the draft *L. bulimoides* genome was incorrect. The mitogenome is usually recovered without much difficulty, as evidenced in metazoan mitogenome skimming methods (Crampton-Platt et al., 2016), as there are usually numerous copies of mitochondrial genomes within a cell for the two copies of autosomal nuclear DNA, providing a natural enrichment for mitochondrial DNA. However, the *L. bulimoides* mitogenome appears to be difficult to assemble from a whole genome sequencing effort (Chapter 3). Similar difficulties were also encountered with the mitogenome recovery efforts for *Calanus* spp., which are also characterized by large and complex genomes (G. Hoarau, personal communication 2021). Additional studies will be needed on the mitochondrial genome of *L. bulimoides* to inform future efforts for assembly and target capture design in other pteropod species.

Reliability of the COI barcode

The mitochondrial cytochrome oxidase I (COI) gene fragment is the most commonly used barcode for species identification and delimitation in animals (Pentinsaari et al., 2016), as it can be reliably amplified with universal primers (Folmer et al., 1994; Geller et al., 2013; Leray et al., 2013). It is also the most frequently used barcoding region for the identification of marine zooplankton species, with ongoing efforts to set up a well-curated reference database for zooplankton (Bucklin et al., 2021; see also https://metazoogene.org). The COI barcoding gene fragment has been shown to resolve population structure in zooplankton species (Burridge et al., 2017b; Goetze et al., 2017; Miyamoto et al., 2012). Across a wide range of taxa, the COI region has also been calibrated as a molecular clock, to allow for estimations of the time of divergence between reproductively isolated populations and species (Hellberg and Vacquier, 1999; Knowlton and Weigt, 1998; Luttikhuizen et al., 2003; Wall-Palmer et al., 2020; Wilke et al., 2009).

In some cases, the COI gene can have peculiarities which cause it to deviate from commonly held expectations of its evolution. For example, the COI gene is usually maternally inherited as a haploid copy. However, in some molluscs, specifically bivalves, doubly uniparental inheritance (i.e., transmission of mitochondria from both parents) can occur, allowing for recombination (Becking et al., 2016; Ghiselli et al., 2021; Hoeh et al., 1991; Zouros et al., 1994). There are also cases of nuclear-mitochondrial fragments (numts) across a wide range of eukaryote clades (Song et al., 2008), where degenerate sequences of the mitochondrial genes are embedded in the nuclear genome. These non-functional sequences accumulate mutations at a much faster rate than the functional mitochondrial copy, potentially leading to overestimations of species diversity. Additionally, selection on the mitochondrial genome and the potential bias caused by introgression can lead to wrong inferences about demography and evolutionary history (Ballard and Whitlock, 2004).

My work further illustrates that while the COI barcode is often used in characterising species diversity, it should not be the only marker used for phylogenetic or phylogeographic studies. In *L. bulimoides*, the patterns of differentiation in the COI gene do not match completely with genome-wide divergence. While I could use COI to delimit between the Atlantic Ocean and Indo-Pacific/Pacific lineages, I was unable to use it to distinguish between the Indo-Pacific and Pacific lineages, which show clear nuclear genome-wide differences (pairwise F_{ST} = 0.300) (Chapter 4). I also recovered only two populations in the Atlantic Ocean based on COI, compared to three populations with genome-wide variation (pairwise F_{ST} = 0.122-0.196) (Chapters 2, 5). Therefore, biologically meaningful information is lost if the analysis relies on the COI gene as the sole marker of divergence. The COI marker is typically expected to accumulate mutations at a faster rate due it its haploid inheritance and faster mutation rates. However, molluscs are predicted to experience a mitochondrial mutation rate similar to the nuclear mutation rate due to their high average nuclear diversity (Allio et al., 2017). In *L. bulimoides*, the COI gene resolved less population structure than the genome-wide nuclear SNPs, despite the exceptionally variable COI fragment, where almost all individuals had a different haplotype (Chapters 2, 4, 5). Resequencing the whole mitochondrial genome would allow us to resolve the evolutionary history of the mitochondrial introgression across the south Atlantic dispersal barrier and complement a genome-wide approach that is needed to obtain the full picture of the evolution into the different lineages seen today.

Assessing barriers to dispersal with genetic and morphometric variation

Shell shape variation in *L. bulimoides* was not always congruent with genetic differentiation. This finding is noteworthy since pteropod shell characters are often used as taxonomic information for delimiting clades and species (Bé and Gilmer, 1977; Burridge et al., 2019; Rampal, 2017; van der Spoel, 1967; van der Spoel and Dadon, 1999). By obtaining standardised orientation photographs of all specimens used for genetic analyses, I was able to compare shell shape variation and genetic divergence across a comprehensive dataset of *L. bulimoides* (Chapters 2, 4, 6). I found that while shell shape differences corresponded with COI haplogroups in the Atlantic AMT22 dataset (Chapter 2), shell shape seemed to vary with sampling location rather than genome-wide differences across broader spatial scales across the Atlantic, Pacific and Indian Oceans (Chapter 4). This discrepancy may be due to the different sets of landmarks used to quantify shell shape (8 landmarks in Chapter 2 and 11 landmarks in Chapter 4). Further investigations are required to determine to what extent shell shape is correlated with genetic differences and/or environmental factors in *L. bulimoides*. In addition, there are likely to be physiological and ecological constraints on shell shape, since shell shape in holoplanktonic gastropods is important for locomotion and predator avoidance (Karakas et al., 2020). In other

gastropods, selection pressures may result in shell shape variation that is incongruent with genetic variation. This could be caused by the maintenance of shell features that directly benefit fitness in the environment despite differing genetic backgrounds (Dowle et al., 2015; Gemmell et al., 2018), or by phenotypic plasticity allowing organisms with similar genetic backgrounds to produce different shell shapes depending on their environment (Hollander et al., 2006; Hollander and Butlin, 2010; Mariani et al., 2012; Vermeij, 2002; Zieritz et al., 2010), or by a combination of both adaptive variation and plasticity.

Besides shell shape, I observed differences in colour of the shell and tissue of *L. bulimoides*, which are correlated with sampling location. Shells with similar colour from one sampling location may indicate that these individuals form part of a swarm which have experienced similar environmental conditions, while sites with specimens of different shell colour may represent individuals from the mixing of swarms exposed to different environmental conditions, given that plankton are often patchy in abundance across their distribution (Levin and Segel, 1976; Omori and Hamner, 1982) and can control their vertical migration (Fossheim and Primicerio, 2008; McLaren, 1963; Trudnowska et al., 2015). Shell pigments are distributed in a phylogenetically relevant manner in most molluscs and the synthesis of colour is energetically costly (Williams, 2017). In the plankton, there is a selective advantage to being transparent to appear invisible in the water column so as to avoid predators (Johnsen, 2001), but this has to be balanced with the protective properties of pigmentation against UV radiation (Bashevkin et al., 2020; Hansson, 2000). It will be interesting to study how these opposing selection pressures play out to result in the colour of the shell and tissue in *L. bulimoides* and other pteropod species. Common garden or reciprocal transplant experiments to investigate the degree of phenotypic plasticity in shell and tissue colour would provide insight into the potential evolutionary mechanisms behind pigmentation in shelled pteropods. However, this remains an unlikely prospect until there are more reliable ways of maintaining pteropods in laboratory culture over their entire life cycle (Howes et al., 2014).

I identified wing spots as a potential distinguishing trait between the sympatric Indo-Pacific and Pacific lineages in the North Pacific Ocean (Chapter 4). These spots occur in pairs on the outer edge of each of the two parapodia (Figure 1A). When the parapodia are retracted, the spots can be visible through the transparent shell as a roughly vertical line of four dark spots (Figure 1B). Because the photos of the specimens were only taken in one standardised (apertural) orientation before destructive DNA extraction, I was not able to conclusively determine if those pigmented spots were absent or not visible on the photographs (Chapter 4). Below, I added additional observations of 29 specimens from a site (KOK1703_03) near Hawai‘i, where both Indo-Pacific and Pacific lineage specimens were identified from the genome-wide dataset. These specimens were photographed both in the apertural orientation, as well as from the opposite side of the shell along its longest

axis, to better be able to observe the presence or absence of the pigmented spots. These specimens can be split into two groups based on their morphological appearance: with pigmented spots (Figure 2A), and without pigmented spots (Figure 2B). By grouping these specimens based on the presence or absence of wing spots, I observed subtle differences in colour between the two groups. Specimens with wing spots tend to have a bright orange tint to their shells and tissues compared to specimens without wing spots, and this may be related to their respective diets and the production of pigment within the snail. Further genetic investigation will be needed to determine if these morphological traits indeed correspond with the split into the Indo-Pacific and Pacific lineages, with the Pacific lineage individuals expected to possess the wing spots (Chapter 4).

For the photographs shown in Figures 1 and 2, and those in Chapter 5, orienting of the shells was done with alcohol-based hand sanitiser. The hand sanitiser allowed for more accurate and efficient positioning of the shells compared to previously used media: fine black sand in Chapters 2 and 4, or white putty in Chapter 4. The easier positioning of the shells in a standardised orientation means that even juveniles, which were previously considered as too challenging to photograph due to their smaller size, can now be included in integrative taxonomy and barcoding pipelines, and provide useful insight into the ontogeny of shelled pteropods. These photographs were made against a black background to highlight shell outlines and tissue colour, which is a standard approach used for scientific publication. However, a white background (as in some specimens for Chapter 4) would be better for the visualisation of shell colour, especially of the aperture, as the mostly transparent aper-

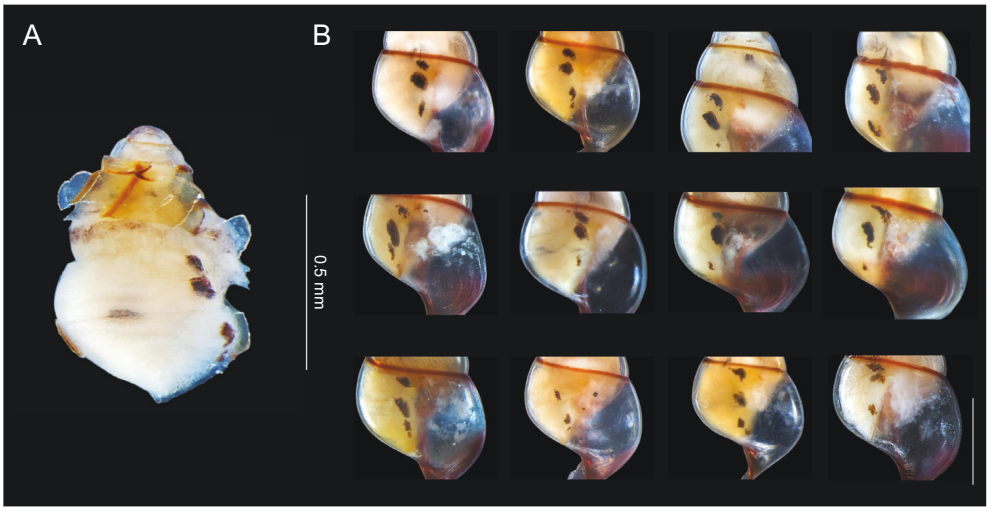


Figure 1 (A) Location of the pigmented spots on the outer edges of the wings of a *Limacina bulimoides* individual, with its shell cracked to expose the tissue, from the sampling location KOK1703_03 off Hawai'i. (B) Close-up of pigmented spots on dorsal view of the 12 individuals in Figure 2A. Scale bar indicates 0.5 mm.

ture blends in against a dark background. With the use of the transparent alcohol gel, the colour of the background can be easily switched, depending on the purpose of the photographs. Hence, this method offers a versatile approach for imaging

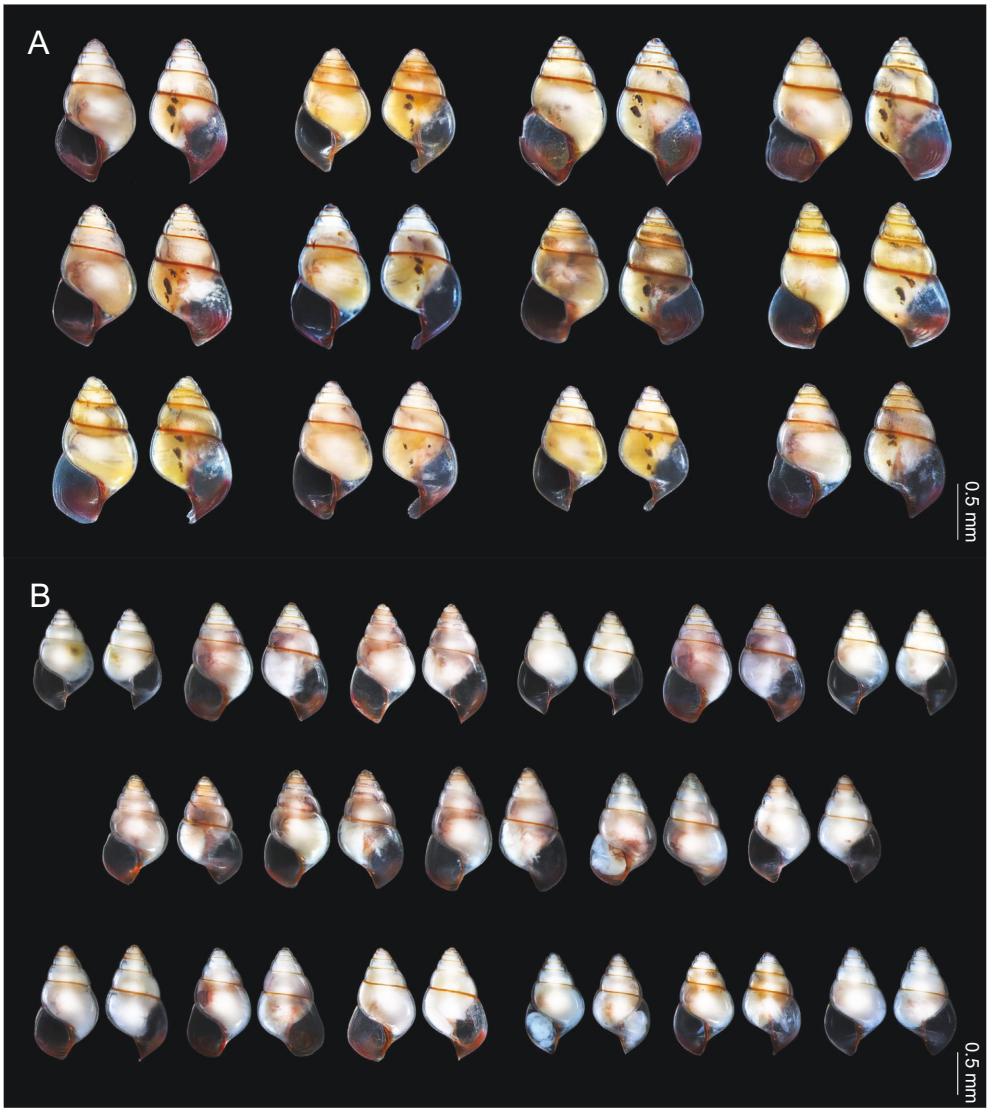


FIGURE 2 Standardised apertural (left) and dorsal (right) views of *Limacina bulimoides* specimens collected from the sampling location KOK1703_03 off Hawai'i, where the Indo-Pacific and Pacific lineages are sympatric. (A) Twelve specimens with wing spots that are visible through their transparent shell on the dorsal view. (B) Seventeen specimens without wing spots. Overall, specimens with wings spots have a brighter orange tint to their tissue and shell compared to those without wing spots, which have paler tissue and shell pigmentation. Scale bar indicates 0.5 mm.

shelled pteropods and potentially other plankton species. A colour standard can also be included in subsequent photographs to ensure that colour remains consistent across all images, so that colour can be quantified objectively for future analyses of shell and tissue colour in *L. bulimoides*.

Towards identifying the nature of dispersal barriers in the Atlantic Ocean

I identified two dispersal barriers in *L. bulimoides* at 14-15°N and 15-18°S in the Atlantic (Chapters 2, 5) that appear to be temporally stable, since they have been recovered over two transects separated by two years and adjacent populations show no signs of recent introgression, hybridisation or gene flow. Broadly similar geographic dispersal barriers for other zooplankton across the equatorial region (Andrews et al., 2014; Burridge et al., 2015; Casteleyn et al., 2010; Norton and Goetze, 2013; Wall-Palmer et al., 2016b) and southern subtropical gyre (Judkins, 2014) have previously been identified. Zooplankton population structure has been attributed to regions of suboptimal habitat leading to low population abundances and inability to sustain a stable population across the unsuitable habitat (Goetze et al., 2017; Norris, 2000; Peijnenburg et al., 2004). Population structure in zooplankton has also been shown to persist over time with sampling across repeated transects (Goetze et al., 2017; Iacchei et al., 2017). It is still unknown whether the distribution or abundance of populations varies in accordance with seasonal environmental fluctuations, and repeated sampling of the Atlantic transect, over different seasons, will be needed to understand the effect of seasonal variability on the population barriers. Additional studies of population genetic structure in other species from the pelagic community can be used to detect if such barriers are species specific, or form multispecies biogeographic breaks (Avise, 2009; Barber et al., 2000; Johannesson et al., 2020; Stanley et al., 2018).

These putative dispersal barriers are surprisingly narrow given the large spatial range of *L. bulimoides* and the seasonally variable and fluid marine environment. This can be attributed to potential reproductive incompatibilities between the Atlantic populations due to historical divergence followed by secondary contact (elaborated upon in the next section). However, strong selection against hybrids or immigrants may also be sufficient to maintain this pattern of genetic structuring (Prada and Hellberg, 2014; Singhal and Moritz, 2012). For this to occur, there should be steep ecological clines, across which selection pressures against the different variants can act. Seascape genomics approaches, which study the association of genetic polymorphisms with particular environmental conditions (Riginos et al., 2016) can be used to identify the physical-chemical parameters that are responsible for the spatial partitioning of *L. bulimoides* populations in the pelagic environment, such as ocean currents, temperature, salinity, carbonate chemistry or food availability. When plotted against the chlorophyll *a* profiles of the AMT22 and AMT24 cruise tracks, the North and South Atlantic populations are located

within oligotrophic subtropical gyre systems, while the Equatorial population is mostly restricted to the more mesotrophic equatorial current system (Chapters 2 and 5). The transitional zones between these habitats may function as barriers mediated by ocean currents, which may lead to limited exchange between the different currents where individuals passing through may still survive but fail to lead to viable offspring reliably. We will need to screen more individuals, including both juveniles and adults, from across these transitional zones to determine whether such advected individuals exist and whether they contribute to gene flow across these dispersal barriers. By including ocean current models to calculate the dispersal distance between sampling locations, and correlating that with the extent of gene flow, it can be determined if dispersal is indeed limiting for the species, despite the commonly held assumption of boundless dispersal in the pelagic environment. This can be informative for predicting future distributions of the species based on models of future climate projections to predict their ability to persist in the face of environmental change. More *in-situ* observations of the ecology of *L. bulimoides* are needed for ecological niche modelling with predictive models of future conditions to identify whether the three Atlantic populations can persist, or whether genetic mixing, facilitated by increased connectivity, will occur due to changes in the ocean environment.

Species boundaries in the open ocean

Historical geographic isolation, followed by secondary contact, can be a possible contributing mechanism to the divergence of *L. bulimoides* in the open ocean, given the geographical separation between lineages in the Atlantic, Pacific and Indian Ocean (Chapter 4). The divergence of the three main lineages, assuming a generation time of one year, was dated to the mid–Pleistocene transition (Chapter 4), a period characterised by global cooling, changing ocean circulation and lengthening of glacial stages (Clark et al., 2006; Elderfield et al., 2012; McClymont et al., 2013). The mid-Pleistocene transition appears to be a driver of evolutionary change, as it has been linked to allopatric speciation in coastal fauna (Yuasa et al., 2021) and pelagic coccolithophores (Filatov et al., 2021) as well as drastic changes in marine biota, such as a mass extinction event of benthic foraminifera coupled with phytoplankton evolution (Kender et al., 2016). The fluctuations in temperature, sea level and ocean currents from glacial-interglacial cycles during the Pleistocene, including an estimated 5 °C decrease in tropical sea surface temperatures during the last glacial maximum (Shakun and Carlson, 2010), could have resulted in repeated opportunities for genetic differentiation and hybridisation of marine plankton (Norris and Hull, 2012; Peijnenburg et al., 2004). Such events can be recovered as genetic signatures of changes in population size (Chapter 4), even in pelagic species known to have high effective population sizes in the modern ocean.

While the ocean is an environment where high dispersal and connectivity is possible, most holoplanktonic species do not manage to colonise the entire ocean

and appear to be locally adapted (Peijnenburg and Goetze, 2013). Ocean currents are considered one of the main driving forces of the spatial structuring of zooplankton (Laso-Jadart et al., 2021; Richter et al., 2019; Sommeria-Klein et al., 2020). Especially for epipelagic plankton, wind-driven surface currents influence the bulk of their transport, with few organisms that survive when advected into the thermohaline circulation of the deep ocean, although rare instances of long distance deep-sea transport have been documented (Banks et al., 2007; Bik et al., 2010; Darling et al., 2000). In addition to ocean currents over large spatial scales, seascape features like atolls, straits and shoals can also influence connectivity and gene flow at local to regional scales (Lal et al., 2017) although these features are probably less relevant to oceanic plankton. Plankton can also be organised into their spatial niches via ecological selection (Ward et al., 2021). Even if locations are connected by currents, not all individuals may survive to reproduce because of exposure to environmental conditions outside their range of tolerance, or competition with the resident species, leading to a reduction in population size and insufficient numbers to sustain a population. Hence, despite the lack of apparent physical barriers, most marine pelagic species can experience geographic isolation, which can facilitate the accumulation of independent genetic divergences on the road to speciation.

Do the genetic differences of the different *L. bulimoides* lineages and populations indicate that they should be classified as different species? This depends on the species concept used. Here I used the reproductive species concept, where species are groups of actually or potentially interbreeding populations that are reproductively isolated from other such groups. Despite genomic evidence pointing towards reproductive isolation between the three main lineages, namely the Atlantic, Indo-Pacific and Pacific lineages, as well as between the three Atlantic populations, I prefer to err on the side of caution and refrain from assigning new species names. By assessing the genomic variability in *L. bulimoides*, I observed that in the Atlantic, Indo-Pacific and Pacific lineages, there was no recent gene flow between any of the lineages (Chapter 4). The three Atlantic populations had a high proportion of ancestry, but also did not show signs of gene flow (Chapter 5). For the absence of gene flow to be maintained over time in a potentially interconnected and fluctuating environment, it seems likely that reproductive isolation plays an important role in separating these lineages and populations. However, I still do not have conclusive data about morphological and ecological differences between the lineages, and thus lack sufficient information for an integrative taxonomy framework to delimit new species (Dayrat, 2005; Padial et al., 2010). It is difficult to recombine species after splitting them, especially in the scientific literature, where naming inconsistencies can confuse and set back progress, hence it is not a decision to be taken lightly to name new species. In addition, there is a matter of how useful the new species boundaries are, when they cannot be identified by eye, and genomic analyses are currently required to distinguish them. However, even with-

out distinct names, the different genetic lineages can and should be recognised as such and be analysed as evolutionarily distinct units. Such units have different evolutionary trajectories and are likely to respond differently to global change stressors such as ocean acidification.

Potential to adapt to future ocean conditions

While additional functional genomic analyses will be required to determine whether *L. bulimoides* has the capacity to adapt to changing conditions, their current high levels of standing genetic variation within lineages and populations across their geographical range is a positive sign (Chapters 4, 5; Bitter et al., 2019; Ørsted et al., 2019; Peijnenburg and Goetze, 2013). The spatial genetic partitioning of this circumglobal species complex and extent of genetic differentiation between populations also suggest that they may be already locally adapted to regional oceanic conditions. Since I used target capture to obtain the genome-wide variation where the probes were designed based on known genomic regions, it could be possible to trace outlier loci and link them to their physiological function. Such loci may be associated with environmental clines, and could be used to identify genes under selection. When shelled pteropods are used as bioindicators to monitor the impact of ocean acidification, it is important to account for differences among their distinct evolutionary lineages, which are likely to have different environmental tolerances. Niche conservatism, the hypothesis where a species' niche remains constant through space and time, cannot be assumed across divergent populations (McGinty et al., 2021). The absence of apparent bottlenecks in the recent demographic history of *L. bulimoides*, as well as current genetic diversity estimates within lineages and populations, suggest that lineages have been relatively unaffected by the repeated periods of glacial-interglacial transitions and increases the likelihood that they may be able to cope with environmental change. The rate of current ocean change, however, is likely to be unprecedented (Brierley and Kingsford, 2009; Zeebe et al., 2016), and it is unknown whether species and communities can keep up with this rate of change (Bell, 2013). Full assessment of the potential of *L. bulimoides* to adapt to future conditions would require an in-depth study into the genetic variation of the functional genes that could provide a selective advantage in a changing ocean, as well as documentation of their rate of evolution.

It is unclear how anthropogenic climate change might affect the current dispersal barriers of *L. bulimoides* in the Atlantic Ocean. The Atlantic Meridional Overturning Circulation has been projected to decline, while gyre systems are expanding polewards (Yang et al., 2020). The distribution ranges of *L. bulimoides* and other subtropical zooplankton species are likely to move polewards as well (Benedetti et al., 2021; Burrows et al., 2011; Poloczanska et al., 2016), assuming the range shifts are within the physiological tolerances of the organisms (Kleypas, 2015). Changes in ocean circulation can severely disrupt oceanic conditions as we know them today, resulting in changes in primary production, and subsequent

trophic mismatches down the food chain (Benedetti et al., 2021; Ji et al., 2010). This can result in shifts in the locations or even the disappearance of the dispersal barriers of *L. bulimoides*. It is also worth noting that increased anthropogenic movement via ballast water from shipping, may play a role in transferring individuals across dispersal barriers, a phenomenon known as cryptic dispersal (David and Loveday, 2018). Over time, with a sufficient number of individuals transferred, genetic homogenisation may occur if there is no selective pressure against migrants, leading to an eventual loss of global genetic diversity.

Overall, calcifying species in tropical and subtropical waters are unlikely to be severely impacted by ocean acidification in the near future, as aragonite saturation levels in the subtropical oceans are projected to decline but still remain well above those needed for calcification during the next several decades (Hartin et al., 2016; Jiang et al., 2015; Orr et al., 2005). Hence, *L. bulimoides* and other (sub)tropical shelled pteropods will not be exposed to very strong selective pressures to maintain their shell integrity, at least not within the next decades. The same cannot be said for the polar and subpolar species of *Limacina*, including *L. helicina* and *L. retroversa* (Magnan et al., 2021), as these species are already facing increasing ocean acidification and temperatures leading to physiological stress (Maas et al., 2018; Mekkes et al., 2021a), regime shifts within the community (Defriez et al., 2016), and are predicted to experience even more drastic rates of change in the near future (Hartin et al., 2016; Terhaar et al., 2020).

Future directions

In summary, I investigated genetic variation across the geographic distribution of the circumglobal shelled pteropod *Limacina bulimoides*, as a first step towards an improved understanding of the adaptive potential of pteropods. The aims of identifying the extent of genetic structuring, evolutionary history and putative mode of speciation in the open ocean were fulfilled, and a preliminary assessment of the adaptability of populations to future ocean changes was made.

In terms of future directions, we can analyse outlier loci and their putative gene functions based on the current genome-wide probes, to gain further insight into the adaptive significance of the observed genetic variation in shelled pteropods. With the current dataset of genome-wide variants across the Atlantic Ocean (Chapter 5), we can detect signals of local adaptation between populations, through methods such as OutFLANK to identify F_{ST} outliers (Whitlock and Lotterhos, 2015) or the Population Branch Statistic (PBS) which can detect population-specific allele frequency deviations (Choudhury et al., 2014; Yi et al., 2010). Because such outlier loci can be traced back to their target capture probes, we can obtain functional genetic information about them, based on their respective categories such as putative biomineralisation genes or conserved pteropod coding regions in the genome (Chapter 3). These sequences can then be compared against online databases to identify related proteins and possible functional annotations, and associate them with their

possible physiological roles in *L. bulimoides*. Having a whole genome assembly available for pteropods would greatly facilitate the detection of selected regions, as physically linked regions such as genomic islands and structural variants, have been commonly implicated in studies of adaptation and speciation (Feder and Nosil, 2010; Kess et al., 2021; Therkildsen et al., 2019; Wolf and Ellegren, 2017; Yeaman and Whitlock, 2011). By combining population genomics and functional studies, it may be possible to directly link past adaptation to selection pressures, and achieve an understanding of their adaptive potential to future climate change.

I propose more sampling in the contact zones, where distinct lineages and populations meet (Johannesson et al., 2020), to gain more insight into the process of genetic differentiation and speciation within *L. bulimoides* across time and space. Additionally, depth-stratified sampling is needed across more locations in the North Pacific Gyre where the Indo-Pacific and Pacific lineages are sympatric, for a better understanding of the mechanisms involved in their reproductive isolation. A region that is currently severely undersampled is the Indo-Australian Archipelago, which represents the area of highest marine diversity (Lohman et al., 2011; Obura, 2016). Sampling in this area could provide further insight into the distribution and divergence between and within the Indo-Pacific and Pacific lineages. Depth-stratified sampling and metabarcoding of gut contents of *L. bulimoides* will be needed to assess if distinct lineages or populations have different ecological niches, which could provide support for ecological divergence. Morphological differences, such as the wing spots between the two sympatric lineages of *L. bulimoides,* could also provide clues for the mechanisms of reproductive isolation.

More fine-scale sampling of *L. bulimoides* along gradients across the two dispersal barriers in the Atlantic Ocean is also needed to characterise the strength of these barriers, as well as the frequency of individuals exhibiting introgressed mitochondrial signals. To be able to assay a much larger number of individuals, while keeping costs down, an important next step is to design a genotyping panel of several candidate loci from the ~2900 loci that were included in the target capture design (Chapter 3). By applying the genotyping panel to individuals collected from the barrier regions, including both juveniles and adults, across different times of the year, it will be possible to determine whether the dispersal barriers in the Atlantic are porous, leading to expatriates or hybrids, or whether these barriers are difficult to cross, lending support to an allopatric model of speciation. A historical perspective with the pteropod fossil record or environmental DNA from deep-sea sediment, as in (Morard et al., 2017), could be useful as additional source of information about previous distributions, relative abundances and population connectivity of genetic lineages.

This thesis is a first effort to assess the genomic variation and population structure of pteropods, which has revealed important information on their biogeographical distribution, barriers to dispersal, and genetic and morphometric variability. From conventional genetic barcoding methods to novel genome-wide probes,

specifically designed for the circumglobal *L. bulimoides*, I found that this species is composed of distinct, spatially separated lineages, with higher than expected spatial genetic structuring. Moving forward, additional sampling, in-depth outlier analyses, and seascape genomics are needed to gain further insight into the drivers of genetic divergence and the evolutionary future of this genetically diverse and ecologically important species complex. The successful application of genome-wide techniques on the non-model zooplankton species *L. bulimoides* paves the way for the application of such techniques to other pteropod species, which are widely studied and used as bioindicators of ocean acidification. This will provide much needed genomic perspectives on their potential to adapt to ongoing and unprecedented changes in oceanic conditions expected in the Anthropocene.

References

10X Genomics [WWW Document], 2019. URL https://www.10xgenomics.com/ (accessed 11.28.19).

Addamo, A.M., Miller, K.J., Häussermann, V., Taviani, M., Machordom, A., 2020. Global-scale genetic structure of a cosmopolitan cold-water coral species. Aquat. Conserv. Mar. Freshw. Ecosyst. 31, 1–14. https://doi.org/10.1002/aqc.3421

Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., Rokhsar, D.S., 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524, 220–224. https://doi.org/10.1038/nature14668

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 19, 1655–1664. https://doi.org/10.1101/gr.094052.109

Alexander, M.A., Scott, J.D., Friedland, K.D., Mills, K.E., Nye, J.A., Pershing, A.J., Thomas, A.C., 2018. Projected sea surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions of Northern Oceans. Elementa 6, 9. https://doi.org/10.1525/elementa.191

Allendorf, F.W., Hohenlohe, P.A., Luikart, G., 2010. Genomics and the future of conservation genetics. Nat. Rev. Genet. 11, 697–709. https://doi.org/10.1038/nrg2844

Allio, R., Donega, S., Galtier, N., Nabholz, B., 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol. Biol. Evol. 34, 2762–2772. https://doi.org/10.1093/molbev/msx197

Álvarez-Noriega, M., Burgess, S.C., Byers, J.E., Pringle, J.M., Wares, J.P., Marshall, D.J., 2020. Global biogeography of marine dispersal potential. Nat. Ecol. Evol. 4, 1196–1203. https://doi.org/10.1038/s41559-020-1238-y

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., Hohenlohe, P.A., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nat. Rev. Genet. 17, 81–92. https://doi.org/10.1038/nrg.2015.28

Andrews, K.R., Norton, E.L., Fernandez-Silva, I., Portner, E., Goetze, E., 2014. Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. Mol. Ecol. 23, 5462–5479. https://doi.org/10.1111/mec.12950

Apolônio Silva De Oliveira, D., Decraemer, W., Moens, T., Dos Santos, G.A.P., Derycke, S., 2017. Low genetic but high morphological variation over more than 1000 km coastline refutes omnipresence of cryptic diversity in marine nematodes. BMC Evol. Biol. 17, 71. https://doi.org/10.1186/s12862-017-0908-0

Attard, C.R.M., Brauer, C.J., Sandoval-Castillo, J., Faulks, L.K., Unmack, P.J., Gilligan, D.M., Beheregaray, L.B., 2018. Ecological disturbance influences adaptive divergence despite high gene flow in golden perch (*Macquaria ambigua*): Implications for management and resilience to climate change. Mol. Ecol. 27, 196–215. https://doi.org/10.1111/mec.14438

Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V, Altshuler, D., Gabriel, S., DePristo, M.A., 2013. From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinforma. 43, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Avise, J.C., 2009. Phylogeography: Retrospect and prospect. J. Biogeogr. 36, 3–15. https://doi.org/10.1111/ j.1365-2699.2008.02032.x

Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P., Antonelli, A., 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proc. Natl. Acad. Sci. U. S. A. 112, 6110–6115. https://doi.org/10.1073/pnas.1423853112

Bagøien, E., Kiørboe, T., 2005. Blind dating—mate finding in planktonic copepods. I. Tracking the pheromone trail of *Centropages typicus*. Mar. Ecol. Prog. Ser. 300, 105–115.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3, e3376. https://doi.org/10.1371/journal.pone.0003376

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x

Banks, S.C., Piggot, M.P., Williamson, J.E., Bové, U., Holbrook, N.J., Beheregaray, L.B., 2007. Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecology 88, 3055–3064. https://doi.org/10.1890/07-0091.1

Barber, P.H., Palumbi, S.R., Erdmann, M. V., Moosa, M.K., 2000. A marine Wallace's line? Nature 406, 692–693. https://doi.org/10.1038/35021135

Barluenga, M., Stölting, K.N., Salzburger, W., Muschick, M., Meyer, A., 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439, 719–723. https://doi.org/10.1038/nature04325

Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Barrio, A.M., Lamichhaney, S., Fan, G., Rafati, N., Pettersson, M., Zhang, H., Dainat, J., Ekman, D., Höppner, M., Jern, P., Martin, M., Nystedt, B., Liu, X., Chen, W., Liang, X., Shi, C., Fu, Y., Ma, K., Zhan, X., Feng, C., Gustafson, U., Rubin, C.J., Almén, M., Blass, M., Casini, M., Folkvord, A., Laikre, L., Ryman, N., Lee, S.Y., Xu, X., Andersson, L., 2016. The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. Elife 5, e12081. https://doi.org/10.7554/eLife.12081

Barth, J.M.I., Berg, P.R., Jonsson, P.R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., Jakobsen, K.S., Johannesson, K., Jorde, P.E., Knutsen, H., Moksnes, P.O., Star, B., Stenseth, N.C., Svedäng, H., Jentoft, S., André, C., 2017. Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Mol. Ecol. 26, 4452–4466. https://doi.org/10.1111/mec.14207

Barton, A.D., Lozier, M.S., Williams, R.G., 2015. Physical controls of variability in North Atlantic phytoplankton communities. Limnol. Oceanogr. 60, 181–197. https://doi.org/10.1002/lno.10011

Bashevkin, S.M., Christy, J.H., Morgan, S.G., 2020. Costs and compensation in zooplankton pigmentation under countervailing threats of ultraviolet radiation and predation. Oecologia 193, 111–123. https://doi.org/10.1007/s00442-020-04648-2

Bé, A.W.H., Gilmer, R., 1977. A zoogeographic and taxonomic review of Euthecosomatous Pteropoda. Ocean. Micropaleontol. 1, 733–808.

Beaugrand, G., Goberville, E., Luczak, C., Kirby, R.R., 2014. Marine biological shifts and climate. Proc. R. Soc. B Biol. Sci. 281, 20133350. https://doi.org/10.1098/rspb.2013.3350

Beaugrand, G., Luczak, C., Edwards, M., 2009. Rapid biogeographical plankton shifts in the North Atlantic Ocean. Glob. Chang. Biol. 15, 1790–1803. https://doi.org/10.1111/j.1365-2486.2009. 01848.x

Becking, L.E., de Leeuw, C.A., Knegt, B., Maas, D.L., de Voogd, N.J., Abdunnur, Suyatna, I., Peijnenburg, K.T.C.A., 2016. Highly divergent mussel lineages in isolated Indonesian marine lakes. PeerJ 4, e2496. https://doi.org/10.7717/peerj.2496

Bedford, J., Ostle, C., Johns, D.G., Atkinson, A., Best, M., Bresnan, E., Machairopoulou, M., Graves, C.A., Devlin, M., Milligan, A., Pitois, S., Mellor, A., Tett, P., McQuatters-Gollop, A., 2020. Lifeform indicators reveal large-scale shifts in plankton across the North-West European shelf. Glob. Chang. Biol. 26, 3482–3497. https://doi.org/10.1111/gcb.15066

Bednaršek, N., Feely, R.A., Beck, M.W., Glippa, O., Kanerva, M., Engström-Öst, J., 2018. El Niño-related thermal stress coupled with upwelling-related ocean acidification negatively impacts cellular to population-level responses in pteropods along the California current system with implications for increased bioenergetic costs. Front. Mar. Sci. 5, 486. https://doi.org/10.3389/fmars.2018.00486

Bednaršek, N., Feely, R.A., Howes, E.L., Hunt, B.P. V., Kessouri, F., León, P., Lischka, S., Maas, A.E., McLaughlin, K., Nezlin, N.P., Sutula, M., Weisberg, S.B., 2019. Systematic review and meta-analysis toward synthesis of thresholds of ocean acidification impacts on calcifying pteropods and interactions with warming. Front. Mar. Sci. 6, 227. https://doi.org/10.3389/fmars.2019.00227

Bednaršek, N., Feely, R.A., Tolimieri, N., Hermann, A.J., Siedlecki, S.A., Waldbusser, G.G., McElhany, P., Alin, S.R., Klinger, T., Moore-Maley, B., Pörtner, H.O., 2017a. Exposure history determines pteropod vulnerability to ocean acidification along the US West Coast. Sci. Rep. 7, 4526. https://doi.org/10.1038/s41598-017-03934-z

Bednaršek, N., Harvey, C.J., Kaplan, I.C., Feely, R.A., Možina, J., 2016. Pteropods on the edge: Cumulative effects of ocean acidification, warming, and deoxygenation. Prog. Oceanogr. 145, 1–24. https://doi.org/10.1016/j.pocean.2016.04.002

Bednaršek, N., Klinger, T., Harvey, C.J., Weisberg, S., McCabe, R.M., Feely, R.A., Newton, J., Tolimieri, N., 2017b. New ocean, new needs: Application of pteropod shell dissolution as a biological indicator for marine resource management. Ecol. Indic. 76, 240–244. https://doi.org/10.1016/j.ecolind.2017.01.025

Bednaršek, N., Možina, J., Vogt, M., O'Brien, C., Tarling, G.A., 2012a. The global distribution of pteropods and their contribution to carbonate and carbon biomass in the modern ocean. Earth Syst. Sci. Data 4, 167–186. https://doi.org/10.5194/essd-4-167-2012

Bednaršek, N., Tarling, G.A., Bakker, D.C.E., Fielding, S., Cohen, A., Kuzirian, A., Mccorkle, D., Lézé, B., Montagna, R., 2012b. Description and quantification of pteropod shell dissolution: A sensitive bioindicator of ocean acidification. Glob. Chang. Biol. 18, 2378–2388. https://doi.org/10.1111/j. 1365- 2486.2012.02668.x

Bednaršek, N., Tarling, G.A., Bakker, D.C.E., Fielding, S., Jones, E.M., Venables, H.J., Ward, P., Kuzirian, A., Lézé, B., Feely, R.A., Murphy, E.J., 2012c. Extensive dissolution of live pteropods in the Southern Ocean. Nat. Geosci. 5, 881–885. https://doi.org/10.1038/ngeo1635

Bell, G., 2013. Evolutionary rescue and the limits of adaptation. Philos. Trans. R. Soc. B Biol. Sci. 368, 20120080. https://doi.org/10.1098/rstb.2012.0080

Belton, J.-M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., Dekker, J., 2012. Hi–C: A comprehensive technique to capture the conformation of genomes. Methods 58, 268–276. https://doi.org/10.1016/j.ymeth.2012.05.001

Bendif, E.M., Nevado, B., Wong, E.L.Y., Hagino, K., Probert, I., Young, J.R., Rickaby, R.E.M., Filatov, D.A., 2019. Repeated species radiations in the recent evolution of the key marine phytoplankton lineage Gephyrocapsa. Nat. Commun. 10, 4234. https://doi.org/10.1038/s41467-019-12169-7

Benedetti, F., Vogt, M., Elizondo, U.H., Righetti, D., Zimmermann, N.E., Gruber, N., 2021. Major restructuring of marine plankton assemblages under global warming. Nat. Commun. 12, 5226. https://doi.org/10.1038/s41467-021-25385-x

Benestan, L., Quinn, B.K., Maaroufi, H., Laporte, M., Clark, F.K., Greenwood, S.J., Rochette, R., Bernatchez, L., 2016. Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). Mol. Ecol. 25, 5073–5092. https://doi.org/10.1111/mec.13811

Berardi, A., Perinelli, D.R., Merchant, H.A., Bisharat, L., Basheti, I.A., Bonacucina, G., Cespi, M., Palmieri, G.F., 2020. Hand sanitisers amid CoViD-19: A critical review of alcohol-based products on the market and formulation approaches to respond to increasing demand. Int. J. Pharm. 584, 119431. https://doi.org/10.1016/j.ijpharm.2020.119431

Bergan, A.J., Lawson, G.L., Maas, A.E., Wang, Z.A., 2017. The effect of elevated carbon dioxide on the sinking and swimming of the shelled pteropod *Limacina retroversa*. ICES J. Mar. Sci. 74, 1893–1905. https://doi.org/10.1093/icesjms/fsx008

Bernard, K.S., 2006. The role of the euthesocome pteropod, *Limacina retroversa*, in the Polar Frontal Zone, Southern Zone. PhD Thesis. Rhodes University.

Bernard, K.S., Froneman, P.W., 2009. The sub-Antarctic euthecosome pteropod, *Limacina retroversa*: Distribution patterns and trophic role. Deep. Res. Part I Oceanogr. Res. Pap. 56, 582–598. https://doi.org/10.1016/j.dsr.2008.11.007

Bernatchez, L., 2016. On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. J. Fish Biol. 89, 2519–2556. https://doi.org/10.1111/jfb.13145

Bernatchez, S., Xuereb, A., Laporte, M., Benestan, L., Steeves, R., Laflamme, M., Bernatchez, L., Mallet, M.A., 2019. Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. Evol. Appl. 12, 587–609. https://doi.org/10.1111/eva.12741

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., Stadler, P.F., 2013. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69, 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Bertola, L.D., Boehm, J.T., Putman, N.F., Xue, A.T., Robinson, J.D., Harris, S., Baldwin, C.C., Overcast, I., Hickerson, M.J., 2020. Asymmetrical gene flow in five co-distributed syngnathids explained by ocean currents and rafting propensity. Proceedings. Biol. Sci. 287, 20200657. https://doi.org/10.1098/rspb.2020.0657

Bi, K., Linderoth, T., Vanderpool, D., Good, J.M., Nielsen, R., Moritz, C., 2013. Unlocking the vault: Next-generation museum population genomics. Mol. Ecol. 22, 6018–6032. https://doi.org/10.1111/mec.12516

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C., Good, J.M., 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13, 403. https://doi.org/10.1186/1471-2164-13-403

Bierne, N., Bonhomme, F., David, P., 2003. Habitat preference and the marine-speciation paradox. Proc. R. Soc. B Biol. Sci. 270, 1399–1406. https://doi.org/10.1098/rspb.2003.2404

Bik, H.M., Thomas, W.K., Lunt, D.H., Lambshead, P.J.D., 2010. Low endemism, continued deep-shallow interchanges, and evidence for cosmopolitan distributions in free-living marine nematodes (order Enoplida). BMC Evol. Biol. 10, 389. https://doi.org/10.1186/1471-2148-10-389

Bilton, D.T., 2018. *Scotolemon doriae* Pavesi, 1878, a soil-dwelling harvestman new to Britain (Opiliones: Phalangodidae). Arachnology 17, 361–363. https://doi.org/10.13156/arac.2017.17.8.361

Bitter, M.C., Kapsenberg, L., Gattuso, J.P., Pfister, C.A., 2019. Standing genetic variation fuels rapid adaptation to ocean acidification. Nat. Commun. 10, 5821. https://doi.org/10.1038/s41467-019-13767-1

Blaimer, B.B., Lloyd, M.W., Guillory, W.X., Brady, S.G., 2016. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. PLoS One 11, e0161531. https://doi.org/10.1371/journal.pone.0161531

Blanco-Bercial, L., Bucklin, A., 2016. New view of population genetics of zooplankton: RAD-seq analysis reveals population structure of the North Atlantic planktonic copepod *Centropages typicus*. Mol. Ecol. 25, 1566–1580. https://doi.org/10.1111/mec.13581

Bode, M., Laakmann, S., Kaiser, P., Hagen, W., Auel, H., Cornils, A., 2017. Unravelling diversity of deep-sea copepods using integrated morphological and molecular techniques. J. Plankton Res. 39, 600–617. https://doi.org/10.1093/plankt/fbx031

Böer, M., Gannefors, C., Kattner, G., Graeve, M., Hop, H., Falk-Petersen, S., 2005. The Arctic pteropod *Clione limacina*: Seasonal lipid dynamics and life-strategy. Mar. Biol. 147, 707–717. https://doi.org/10.1007/s00227-005-1607-8

Boetzer, M., Henkel, C. V., Jansen, H.J., Butler, D., Pirovano, W., 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27, 578–579. https://doi.org/10.1093/bioinformatics/btq683

Bogan, S.N., Johnson, K.M., Hofmann, G.E., 2020. Changes in genome-wide methylation and gene expression in response to future pCO_2 extremes in the Antarctic pteropod *Limacina helicina antarctica* 6, 788. https://doi.org/10.3389/fmars.2019.00788

Bongaerts, P., Cooke, I.R., Ying, H., Wels, D., den Haan, S., Hernandez-Agreda, A., Brunner, C.A., Dove, S., Englebert, N., Eyal, G., Forêt, S., Grinblat, M., Hay, K.B., Harii, S., Hayward, D.C., Lin, Y., Mihaljević, M., Moya, A., Muir, P., Sinniger, F., Smallhorn-West, P., Torda, G., Ragan, M.A., van Oppen, M.J.H., Hoegh-Guldberg, O., 2021. Morphological stasis masks ecologically divergent coral species on tropical reefs. Curr. Biol. 31, 2286-2298.E8. https://doi.org/10.1016/j.cub.2021.03.028

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., Du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Bowen, B.W., Gaither, M.R., DiBattista, J.D., Iacchei, M., Andrews, K.R., Grant, W.S., Toonen, R.J., Briggs, J.C., 2016. Comparative phylogeography of the ocean planet. Proc. Natl. Acad. Sci. 113, 7962–7969. https://doi.org/10.1073/pnas.1602404113

Bowen, B.W., Rocha, L.A., Toonen, R.J., Karl, S.A., 2013. The origins of tropical marine biodiversity. Trends Ecol. Evol. 28, 359–366. https://doi.org/10.1016/j.tree.2013.01.018

Boyce, J.M., Pittet, D., 2002. Guideline for hand hygiene in health-care settings: Recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. Am. J. Infect. Control 30, S1–S46. https://doi.org/10.1067/mic.2002.130391

Bradbury, I.R., Laurel, B., Snelgrove, P.V.R., Bentzen, P., Campana, S.E., 2008. Global patterns in marine dispersal estimates: The influence of geography, taxonomic category and life history. Proc. R. Soc. B Biol. Sci. 275, 1803–1809. https://doi.org/10.1098/rspb.2008.0216

Bragg, J.G., Potter, S., Bi, K., Moritz, C., 2016. Exon capture phylogenomics: efficacy across scales of divergence. Mol. Ecol. Resour. 16, 1059–1068. https://doi.org/10.1111/1755-0998.12449

Brierley, A.S., Kingsford, M.J., 2009. Impacts of climate change on marine organisms and ecosystems. Curr. Biol. 19, R602–R614. https://doi.org/10.1016/j.cub.2009.05.046

Briscoe, D.K., Hobday, A.J., Carlisle, A., Scales, K., Eveson, J.P., Arrizabalaga, H., Druon, J.N., Fromentin, J.M., 2017. Ecological bridges and barriers in pelagic ecosystems. Deep. Res. Part II Top. Stud. Oceanogr. 140, 182–192. https://doi.org/10.1016/j.dsr2.2016.11.004

Broad Institute, 2009. *Aplysia* Genome Project [WWW Document]. URL https://www.broadinstitute.org/aplysia/aplysia-genome-project (accessed 11.28.19).

Broad Institute, 2019. Picard [WWW Document]. URL http://broadinstitute.github.io/picard/ (accessed 11.28.19).

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., Roychoudhury, A., 2012. Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29, 1917–1932. https://doi.org/10.1093/molbev/mss086

Bryndum-Buchholz, A., Tittensor, D.P., Blanchard, J.L., Cheung, W.W.L., Coll, M., Galbraith, E.D., Jennings, S., Maury, O., Lotze, H.K., 2019. Twenty-first-century climate change impacts on marine animal biomass and ecosystem structure across ocean basins. Glob. Chang. Biol. 25, 459–472. https://doi.org/10.1111/gcb.14512

Bucklin, A., DiVito, K.R., Smolina, I., Choquet, M., Questel, J.M., Hoarau, G., O'Neill, R.J., 2018. Population Genomics of Marine Zooplankton, in: Oleksiak, M., Rajora, O.P. (Eds.), Population Genomics: Marine Organisms. Springer International Publishing, Cham, Switzerland, pp. 61–102. https://doi.org/10.1007/13836_2017_9

Bucklin, A., Lindeque, P.K., Rodriguez-Ezpeleta, N., Albaina, A., Lehtiniemi, M., 2016. Metabarcoding of marine zooplankton: Prospects, progress and pitfalls. J. Plankton Res. 38, 393–400. https://doi.org/10.1093/plankt/fbw023

Bucklin, A., Peijnenburg, K.T.C.A., Kosobokova, K., Machida, R.J., 2021a. New insights into biodiversity, biogeography, ecology, and evolution of marine zooplankton based on molecular approaches. ICES J. Mar. Sci. 78, 3281–3287. https://doi.org/10.1093/icesjms/fsab198

Bucklin, A., Peijnenburg, K.T.C.A., Kosobokova, K.N., O'Brien, T.D., Blanco-B, L., Cornils, A., Falkenhaug, T., Hopcroft, R.R., Hosia, A., Laakmann, S., Chaolun, L., Martell, L., Questel, J.M., Wall-Palmer, D., Wang, M., Wiebe, P.H., Weydmann-Zwolicka, A., O'Brien, T.D., Blanco-Bercial, L., Cornils, A., Falkenhaug, T., Hopcroft, R.R., Hosia, A., Laakmann, S., Li, C., Martell, L., Questel, J.M., Wall-Palmer, D., Wang, M., Wiebe, P.H., Weydmann-Zwolicka, A., 2021b. Toward a global reference database of COI barcodes for marine zooplankton. Mar. Biol. 168, 1–26. https://doi.org/10.1007/s00227-021-03887-y

Buitenhuis, E.T., Le Quéré, C., Bednaršek, N., Schiebel, R., 2019. Large contribution of pteropods to shallow $CaCO_3$ export. Global Biogeochem. Cycles 33, 458–468. https://doi.org/10.1029/2018GB006110

Burridge, A.K., Goetze, E., Raes, N., Huisman, J., Peijnenburg, K.T.C.A., 2015. Global biogeography and evolution of Cuvierina pteropods. BMC Evol. Biol. 15, 39. https://doi.org/10.1186/s12862-015-0310-8

Burridge, A.K., Goetze, E., Wall-Palmer, D., Le Double, S.L., Huisman, J., Peijnenburg, K.T.C.A., 2017a. Diversity and abundance of pteropods and heteropods along a latitudinal gradient across the Atlantic Ocean. Prog. Oceanogr. 158, 213–223. https://doi.org/10.1016/j.pocean.2016.10.001

Burridge, A.K., Hörnlein, C., Janssen, A.W., Hughes, M., Bush, S.L., Marlétaz, F., Gasca, R., Pierrot-Bults, A.C., Michel, E., Todd, J.A., Young, J.R., Osborn, K.J., Menken, S.B.J., Peijnenburg, K.T.C.A., 2017b. Time-calibrated molecular phylogeny of pteropods. PLoS One 12, e0177325. https://doi.org/10.1371/journal.pone.0177325

Burridge, A.K., Tump, M., Vonk, R., Goetze, E., Peijnenburg, K.T.C.A., 2017c. Diversity and distribution of hyperiid amphipods along a latitudinal transect in the Atlantic Ocean. Prog. Oceanogr. 158, 224–235. https://doi.org/10.1016/j.pocean.2016.08.003

Burridge, A.K., Van Der Hulst, R., Goetze, E., Peijnenburg, K.T.C.A., 2019. Assessing species boundaries in the open sea: an integrative taxonomic approach to the pteropod genus *Diacavolinia*. Zool. J. Linn. Soc. 187, 1016–1040. https://doi.org/10.1093/zoolinnean/zlz049

Burrows, M.T., Schoeman, D.S., Buckley, L.B., Moore, P., Poloczanska, E.S., Brander, K.M., Brown, C., Bruno, J.F., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C. V., Kiessling, W., O'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F.B., Sydeman, W.J., Richardson, A.J., 2011. The pace of shifting climate in marine and terrestrial ecosystems. Science 334, 652–655. https://doi.org/10.1126/science.1210288

Busch, D.S., Maher, M., Thibodeaub, P., McElhany, P., 2014. Shell condition and survival of Puget Sound pteropods are impaired by ocean acidification conditions. PLoS One 9, e105884. https://doi.org/10.1371/journal.pone.0105884

Bush, G.L., 1966. The taxonomy, cytology and evolution of the genus *Rhagoletis* in North America (Diptera, Tephritidae). Bull. Mus. Comp. Zool. 134, 431–562.

Butlin, R.K., Galindo, J., Grahame, J.W., 2008. Review. Sympatric, parapatric or allopatric: The most important way to classify speciation? Philos. Trans. R. Soc. B Biol. Sci. 363, 2997–3007. https://doi.org/10.1098/rstb.2008.0076

Butlin, R.K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., Coyne, J.A., Galindo, J., Grahame, J.W., Hollander, J., Kemppainen, P., Martínez-Fernández, M., Panova, M., Quesada, H., Johannesson, K., Rolán-Alvarez, E., 2014. Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. Evolution 68, 935–949. https://doi.org/10.1111/evo.12329

Butlin, R.K., Stankowski, S., 2020. Is it time to abandon the biological species concept? No. Natl. Sci. Rev. 7, 1400–1401. https://doi.org/10.1093/nsr/nwaa109

Caley, T., Giraudeau, J., Malaizé, B., Rossignol, L., Pierre, C., 2012. Agulhas leakage as a key process in the modes of Quaternary climate changes. Proc. Natl. Acad. Sci. U. S. A. 109, 6835–6839. https://doi.org/10.1073/pnas.1115545109

Callomon, P., 2019. Standard views for imaging mollusk shells. Am. Malacol. Soc.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421. https://doi.org/10.1186/1471-2105-10-421

Cartagena-Sierra, A., Berke, M.A., Robinson, R.S., Marcks, B., Castañeda, I.S., Starr, A., Hall, I.R., Hemming, S.R., LeVay, L.J., Party, E. 361 S., 2021. Latitudinal migrations of the Subtropical Front at the Agulhas Plateau through the Mid-Pleistocene Transition. Paleoceanogr. Paleoclimatology 36, e2020PA004084. https://doi.org/10.1029/2020pa004084

Casteleyn, G., Leliaert, F., Backeljau, T., Debeer, A.E., Kotaki, Y., Rhodes, L., Lundholm, N., Sabbe, K., Vyverman, W., 2010. Limits to gene flow in a cosmopolitan marine planktonic diatom. Proc. Natl. Acad. Sci. U. S. A. 107, 12952–12957. https://doi.org/10.1073/pnas.1001380107

Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7. https://doi.org/10.1186/s13742-015-0047-8

Chaudhary, C., Richardson, A.J., Schoeman, D.S., Costello, M.J., 2021. Global warming is causing a more pronounced dip in marine species richness around the equator. Proc. Natl. Acad. Sci. U. S. A. 118, e2015094118. https://doi.org/10.1073/pnas.2015094118

Chevin, L.M., Gallet, R., Gomulkiewicz, R., Holt, R.D., Fellous, S., 2013. Phenotypic plasticity in evolutionary rescue experiments. Philos. Trans. R. Soc. B Biol. Sci. 368, 20120089. https://doi.org/10.1098/rstb.2012.0089

Chivers, W.J., Walne, A.W., Hays, G.C., 2017. Mismatch between marine plankton range movements and the velocity of climate change. Nat. Commun. 8, 14434. https://doi.org/10.1038/ncomms14434

Choo, L.Q., Bal, T.M.P., Choquet, M., Smolina, I., Ramos-Silva, P., Marlétaz, F., Kopp, M., Hoarau, G., Peijnenburg, K.T.C.A., 2020. Novel genomic resources for shelled pteropods: A draft genome and target capture probes for *Limacina bulimoides*, tested for cross-species relevance. BMC Genomics 21, 11. https://doi.org/10.1186/s12864-019-6372-z

Choo, L.Q., Bal, T.M.P., Goetze, E., Peijnenburg, K.T.C.A., 2021. Oceanic dispersal barriers in a holoplanktonic gastropod. J. Evol. Biol. 34, 224–240. https://doi.org/10.1111/jeb.13735

Choquet, M., 2017. Combining ecological and molecular approaches to redefine the baseline knowledge of the genus *Calanus* in the North Atlantic and the Arctic Oceans (PhD thesis).

Choquet, M., Kosobokova, K., Kwaśniewski, S., Hatlebakk, M., Dhanasiri, A.K.S., Melle, W., Daase, M., Svensen, C., Søreide, J.E., Hoarau, G., 2018a. Can morphology reliably distinguish between the copepods *Calanus finmarchicus* and *C. glacialis*, or is DNA the only way? Limnol. Oceanogr. Methods 16, 237–252. https://doi.org/10.1002/lom3.10240

Choquet, M., Smolina, I., Dhanasiri, A.K.S., Kopp, M., Jueterbock, A., Sundaram, A.Y.M., Hoarau, G., 2018b. Towards population genomics in non-model species with large genomes; a case study of the marine zooplankton *Calanus finmarchicus*. R. Soc. Open Sci. 6, 180608. https://doi.org/10.1098/RSOS.180608

Choquet, M., Smolina, I., Dhanasiri, A.K.S., Kopp, M., Jueterbock, A., Sundaram, A.Y.M., Hoarau, G., 2019. Towards population genomics in non-model species with large genomes: a case study of the marine zooplankton *Calanus finmarchicus*. R. Soc. Open Sci. 6, 1–36. https://doi.org/10.1098/RSOS.180608

Choudhury, A., Hazelhurst, S., Meintjes, A., Achinike-Oduaran, O., Aron, S., Gamieldien, J., Jalali Sefid Dashti, M., Mulder, N., Tiffin, N., Ramsay, M., 2014. Population-specific common SNPs reflect demographic histories and highlight regions of genomic plasticity with functional relevance. BMC Genomics 15, 437. https://doi.org/10.1186/1471-2164-15-437

Chung, J., Son, D.-S., Jeon, H.-J., Kim, K.-M., Park, G., Ryu, G.H., Park, W.-Y., Park, D., 2016. The minimal amount of starting DNA for Agilent's hybrid capture-based targeted massively parallel sequencing. Sci. Rep. 6, 26732. https://doi.org/10.1038/srep26732

Clark, P.U., Archer, D., Pollard, D., Blum, J.D., Rial, J.A., Brovkin, V., Mix, A.C., Pisias, N.G., Roy, M., 2006. The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO_2. Quat. Sci. Rev. 25, 3150–3184. https://doi.org/10.1016/j.quascirev.2006.07.008

Cohen, J.H., Forward Jr, R.B., 2016. Zooplankton diel vertical migration – A review of proximate control, in: Oceanography and Marine Biology. CRC Press, pp. 89–122.

Colliard, C., Sicilia, A., Turrisi, G.F., Arculeo, M., Perrin, N., Stöck, M., 2010. Strong reproductive barriers in a narrow hybrid zone of west-Mediterranean green toads (*Bufo viridis* subgroup) with Plio-Pleistocene divergence. BMC Evol. Biol. 10, 232. https://doi.org/10.1186/1471-2148-10-232

Comeau, S., Gattuso, J.P., Nisumaa, A.M., Orr, J., 2012. Impact of aragonite saturation state changes on migratory pteropods. Proc. R. Soc. B Biol. Sci. 279, 732–738. https://doi.org/10.1098/rspb.2011.0910

Conde-Padín, P., Carvajal-Rodríguez, A., Carballo, M., Caballero, A., Rolán-Alvarez, E., 2007. Genetic variation for shell traits in a direct-developing marine snail involved in a putative sympatric ecological speciation process. Evol. Ecol. 21, 635–650. https://doi.org/10.1007/s10682-006-9142-8

Conley, K.R., Lombard, F., Sutherland, K.R., 2018. Mammoth grazers on the ocean's minuteness: a review of selective feeding using mucous meshes. Proc. R. Soc. B Biol. Sci. 285, 20180056. https://doi.org/10.1098/rspb.2018.0056

Cornils, A., 2015. Non-destructive DNA extraction for small pelagic copepods to perform integrative taxonomy. J. Plankton Res. 37, 6–10. https://doi.org/10.1093/plankt/fbu105

Cornils, A., Wend-Heckmann, B., Held, C., 2017. Global phylogeography of *Oithona similis* s.l. (Crustacea, Copepoda, Oithonidae) – A cosmopolitan plankton species or a complex of cryptic lineages? Mol. Phylogenet. Evol. 107, 473–485. https://doi.org/10.1016/j.ympev.2016.12.019

Corse, E., Rampal, J., Cuoc, C., Pech, N., Perez, Y., Gilles, A., 2013. Phylogenetic analysis of Thecosomata Blainville, 1824 (Holoplanktonic Opisthobranchia) using morphological and molecular data. PLoS One 8, e59439. https://doi.org/10.1371/journal.pone.0059439

Costello, M.J., Tsai, P., Wong, P.S., Cheung, A.K.L., Basher, Z., Chaudhary, C., 2017. Marine biogeographic realms and species endemicity. Nat. Commun. 8, 1057. https://doi.org/10.1038/s41467-017-01121-2

Cowman, P.F., Bellwood, D.R., 2013. Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. Proc. R. Soc. B Biol. Sci. 280, 20131541. https://doi.org/10.1098/rspb.2013.1541

Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates, Sunderland, MA.

Crampton-Platt, A., Yu, D.W., Zhou, X., Vogler, A.P., 2016. Mitochondrial metagenomics: letting the genes out of the bottle. Gigascience 5, 15. https://doi.org/10.1186/s13742-016-0120-y

Crow, K.D., Munehara, H., Kanamoto, Z., Balanov, A., Antonenko, D., Bernardi, G., 2007. Maintenance of species boundaries despite rampant hybridization between three species of reef fishes (Hexagrammidae): Implications for the role of selection. Biol. J. Linn. Soc. 91, 135–147. https://doi.org/10.1111/j.1095-8312.2007.00786.x

Cruz, R.A.L., Pante, M.J.R., Rohlf, F.J., 2012. Geometric morphometric analysis of shell shape variation in *Conus* (Gastropoda: Conidae). Zool. J. Linn. Soc. 165, 296–310. https://doi.org/10.1111/j.1096-3642.2011.00806.x

Dadon, J., 1990. Annual Cycle of *Limacina retroversa* in Patagonian Waters. Am. Malacol. Bull. 8, 77–84.

Dam, H.G., 2013. Evolutionary adaptation of marine zooplankton to global change. Ann. Rev. Mar. Sci. 5, 349–370. https://doi.org/10.1146/annurev-marine-121211-172229

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., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Darling, K.F., Wade, C.M., Stewart, I.A., Kroon, D., Dingle, R., Leigh Brown, A.J., 2000. Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. Nature 405, 43–47. https://doi.org/10.1038/35011002

Darwin, C., 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London, UK.

David, A.A., Loveday, B.R., 2018. The role of cryptic dispersal in shaping connectivity patterns of marine populations in a changing world. J. Mar. Biol. Assoc. United Kingdom 98, 647–655. https://doi.org/10.1017/S0025315417000236

Davis, K.E., Hill, J., Astrop, T.I., Wills, M.A., 2016. Global cooling as a driver of diversification in a major marine clade. Nat. Commun. 7, 13003. https://doi.org/10.1038/ncomms13003

Dayrat, B., 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85, 407–415. https://doi.org/10.1111/j.1095-8312.2005.00503.x

Dayrat, B., Tillier, A., Lecointre, G., Tillier, S., 2001. New clades of euthyneuran gastropods (Mollusca) from 28S rRNA sequences. Mol. Phylogenet. Evol. 19, 225–235. https://doi.org/10.1006/mpev.2001.0926

De Vargas, C., Norris, R., Zaninetti, L., Gibb, S.W., Pawlowski, J., 1999. Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. Proc. Natl. Acad. Sci. U. S. A. 96, 2864–2868. https://doi.org/10.1073/pnas.96.6.2864

De Wit, P., Pespeni, M.H., Palumbi, S.R., 2015. SNP genotyping and population genomics from expressed sequences – current advances and future possibilities. Mol. Ecol. 24, 2310–2323. https://doi.org/10.1111/mec.13165

Deagle, B.E., Faux, C., Kawaguchi, S., Meyer, B., Jarman, S.N., 2015. Antarctic krill population genomics: Apparent panmixia, but genome complexity and large population size muddy the water. Mol. Ecol. 24, 4943–4959. https://doi.org/10.1111/mec.13370

Defriez, E.J., Sheppard, L.W., Reid, P.C., Reuman, D.C., 2016. Climate change-related regime shifts have altered spatial synchrony of plankton dynamics in the North Sea. Glob. Chang. Biol. 22, 2069–2080. https://doi.org/10.1111/gcb.13229

Depristo, M.A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J.R., Hartl, C., Philippakis, A.A., Del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–501. https://doi.org/10.1038/ng.806

Després, L., 2019. One, two or more species? Mitonuclear discordance and species delimitation. Mol. Ecol. 28, 3845–3847. https://doi.org/10.1111/mec.15211

Donelson, J.M., Sunday, J.M., Figueira, W.F., Gaitán-Espitia, J.D., Hobday, A.J., Johnson, C.R., Leis, J.M., Ling, S.D., Marshall, D., Pandolfi, J.M., Pecl, G., Rodgers, G.G., Booth, D.J., Munday, P.L., 2019. Understanding interactions between plasticity, adaptation and range shifts in response to marine environmental change. Philos. Trans. R. Soc. B Biol. Sci. 374, 20180186. https://doi.org/10.1098/rstb.2018.0186

Doney, S.C., Busch, D.S., Cooley, S.R., Kroeker, K.J., 2020. The impacts of ocean acidification on marine ecosystems and reliant human communities. Annu. Rev. Environ. Resour. 45, 83–112. https://doi.org/10.1146/annurev-environ-012320-083019

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N., Sydeman, W.J., Talley, L.D., 2012. Climate change impacts on marine ecosystems. Ann. Rev. Mar. Sci. 4, 11–37. https://doi.org/10.1146/annurev-marine-041911-111611

Dowle, E.J., Morgan-Richards, M., Brescia, F., Trewick, S.A., 2015. Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. Mol. Ecol. 24, 4205–4221. https://doi.org/10.1111/mec.13302

Drouin, K.L., Lozier, M.S., Johns, W.E., 2021. Variability and Trends of the South Atlantic Subtropical Gyre. J. Geophys. Res. Ocean. 126, e2020JC016405. https://doi.org/10.1029/2020JC016405

Edmonds, S.L., MacInga, D.R., Mays-Suko, P., Duley, C., Rutter, J., Jarvis, W.R., Arbogast, J.W., 2012. Comparative efficacy of commercially available alcohol-based hand rubs and World Health Organization-recommended hand rubs: Formulation matters. Am. J. Infect. Control 40, 521–525. https://doi.org/10.1016/j.ajic.2011.08.016

Eizaguirre, C., Baltazar-Soares, M., 2014. Evolutionary conservation-evaluating the adaptive potential of species. Evol. Appl. 7, 963–967. https://doi.org/10.1111/eva.12227

Ekblom, R., Galindo, J., 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107, 1–15. https://doi.org/10.1038/hdy.2010.152

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I.N., Hodell, D., Piotrowski, A.M., 2012. Evolution of ocean temperature and ice volume through the mid-pleistocene climate transition. Science 337, 704–709. https://doi.org/10.1126/science.1221294

Ellingson, R.A., Krug, P.J., 2015. Reduced genetic diversity and increased reproductive isolation follow population-level loss of larval dispersal in a marine gastropod. Evolution 70, 18–37. https://doi.org/10.1111/evo.12830

Endler, J.A., 1977. Geographic variation, speciation, and clines. Princeton University Press.

Etter, P.D., Preston, J.L., Bassham, S., Cresko, W.A., Johnson, E.A., 2011. Local De Novo Assembly of RAD Paired-End Contigs Using Short Sequencing Reads. PLoS One 6, e18561. https://doi.org/10.1371/journal.pone.0018561

Ewing, G.B., Jensen, J.D., 2016. The consequences of not accounting for background selection in demographic inference. Mol. Ecol. 25, 135–141. https://doi.org/10.1111/mec.13390

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. Mol. Ecol. Resour. 10, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fabry, V., McClintock, J., Mathis, J., Grebmeier, J., 2009. Ocean acidification at high latitudes: The bellwether. Oceanography 22, 160–171. https://doi.org/10.5670/oceanog.2009.105

Faria, R., Johannesson, K., Stankowski, S., 2021. Speciation in marine environments: Diving under the surface. J. Evol. Biol. 34, 4–15. https://doi.org/10.1111/jeb.13756

Fauvelot, C., Zuccon, D., Borsa, P., Grulois, D., Magalon, H., Riquet, F., Andréfouët, S., Berumen, M.L., Sinclair-Taylor, T.H., Gélin, P., Behivoke, F., Poorten, J.J., Strong, E.E., Bouchet, P., 2020. Phylogeographical patterns and a cryptic species provide new insights into Western Indian Ocean giant clams phylogenetic relationships and colonization history. J. Biogeogr. 47, 1086–1105. https://doi.org/10.1111/jbi.13797

Feder, J.L., Nosil, P., 2010. The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. Evolution 64, 1729–1747. https://doi.org/10.1111/j.1558-5646.2009.00943.x

Feely, R.A., Doney, S.C., Cooley, S.C., 2009. Ocean acidification: Present conditions and future changes in a high-CO_2 world. Oceanography 22, 36–47.

Feely, R.A., Sabine, C.L., Lee, K., Berelson, W., Kleypas, J., Fabry, V.J., Millero, F.J., 2004. Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. Science 305, 362–366. https://doi.org/10.1126/science.1097329

Filatov, D.A., Bendif, E.M., Archontikis, O.A., Hagino, K., Rickaby, R.E.M., 2021. The mode of speciation during a recent radiation in open-ocean phytoplankton. Curr. Biol. 31, 1–11. https://doi.org/10.1016/j.cub.2021.09.073

Fišer, C., Robinson, C.T., Malard, F., 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27, 613–635. https://doi.org/10.1111/mec.14486

Fitzpatrick, B.M., Fordyce, J.A., Gavrilets, S., 2009. Pattern, process and geographic modes of speciation. J. Evol. Biol. 22, 2342–2347. https://doi.org/10.1111/j.1420-9101.2009.01833.x

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299. https://doi.org/10.1071/ZO9660275

Foote, A.D., Vijay, N., Ávila-Arcos, M.C., Baird, R.W., Durban, J.W., Fumagalli, M., Gibbs, R.A., Hanson, M.B., Korneliussen, T.S., Martin, M.D., Robertson, K.M., Sousa, V.C., Vieira, F.G., Vinař, T., Wade, P., Worley, K.C., Excoffier, L., Morin, P.A., Gilbert, M.T.P., Wolf, J.B.W., 2016. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. Nat. Commun. 7, 11693. https://doi.org/10.1038/ncomms11693

Förster, D.W., Bull, J.K., Lenz, D., Autenrieth, M., Paijmans, J.L.A., Kraus, R.H.S., Nowak, C., Bayerl, H., Kuehn, R., Saveljev, A.P., Sindičić, M., Hofreiter, M., Schmidt, K., Fickel, J., 2018. Targeted resequencing of coding DNA sequences for SNP discovery in nonmodel species. Mol. Ecol. Resour. 18, 1356–1373. https://doi.org/10.1111/1755-0998.12924

Fossheim, M., Primicerio, R., 2008. Habitat choice by marine zooplankton in a high-latitude ecosystem. Mar. Ecol. Prog. Ser. 364, 47–56. https://doi.org/10.3354/meps07483

Fossheim, M., Primicerio, R., Johannesen, E., Ingvaldsen, R.B., Aschan, M.M., Dolgov, A. V., 2015. Recent warming leads to a rapid borealization of fish communities in the Arctic. Nat. Clim. Chang. 5, 673–677. https://doi.org/10.1038/nclimate2647

Foster, B.A., 1987. Composition and abundance of zooplankton under the spring sea-ice of McMurdo Sound, Antarctica. Polar Biol. 8, 41–48. https://doi.org/10.1007/BF00297163

Fourdrilis, S., Mardulyn, P., Hardy, O.J., Jordaens, K., de Frias Martins, A.M., Backeljau, T., 2016. Mitochondrial DNA hyperdiversity and its potential causes in the marine periwinkle *Melarhaphe neritoides* (Mollusca: Gastropoda). PeerJ 4, e2549. https://doi.org/10.7717/peerj.2549

Franco, B.C., Defeo, O., Piola, A.R., Barreiro, M., Yang, H., Ortega, L., Gianelli, I., Castello, J.P., Vera, C., Buratti, C., Pájaro, M., Pezzi, L.P., Möller, O.O., 2020. Climate change impacts on the atmospheric circulation, ocean, and fisheries in the southwest South Atlantic Ocean: a review. Clim. Change 162, 2359–2377. https://doi.org/10.1007/s10584-020-02783-6

Freeman, B.G., Pennell, M.W., 2021. The latitudinal taxonomy gradient. Trends Ecol. Evol. 36, 778–786. https://doi.org/10.1016/j.tree.2021.05.003

Friesen, V.L., Smith, A.L., Gómez-Díaz, E., Bolton, M., Furness, R.W., González-Solís, J., Monteiro, L.R., 2007. Sympatric speciation by allochrony in a seabird. Proc. Natl. Acad. Sci. U. S. A. 104, 18589–18594. https://doi.org/10.1073/pnas.0700446104

Gaggiotti, O.E., Bekkevold, D., Jørgensen, H.B.H., Foll, M., Carvalho, G.R., Andre, C., Ruzzante, D.E., 2009. Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. Evol. Int. J. Org. Evol. 63, 2939–2951. https://doi.org/10.1111/j.1558-5646.2009.00779.x

Gagnaire, P.A., 2020. Comparative genomics approach to evolutionary process connectivity. Evol. Appl. 13, 1320–1334. https://doi.org/10.1111/eva.12978

Gagnaire, P.A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., Bierne, N., 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. Evol. Appl. 8, 769–786. https://doi.org/10.1111/eva.12288

Galtier, N., 2016. Adaptive protein evolution in animals and the effective population size hypothesis. PLOS Genet. 12, e1005774. https://doi.org/10.1371/journal.pgen.1005774

Garciá Molinos, J., Halpern, B.S., Schoeman, D.S., Brown, C.J., Kiessling, W., Moore, P.J., Pandolfi, J.M., Poloczanska, E.S., Richardson, A.J., Burrows, M.T., 2016. Climate velocity and the future global redistribution of marine biodiversity. Nat. Clim. Chang. 6, 83–88. https://doi.org/10.1038/nclimate2769

Gardner, J., Manno, C., Bakker, D.C.E., Peck, V.L., Tarling, G.A., 2018. Southern Ocean pteropods at risk from ocean warming and acidification. Mar. Biol. 165, 8. https://doi.org/10.1007/s00227-017-3261-3

Gattuso, J.-P., Magnan, A., Billé, R., Cheung, W.W.L., Howes, E.L., Joos, F., Allemand, D., Bopp, L., Cooley, S.R., Eakin, C.M., Hoegh-Guldberg, O., Kelly, R.P., Pörtner, H.-O., Rogers, A.D., Baxter, J.M., Laffoley, D., Osborn, D., Rankovic, A., Rochette, J., Sumaila, U.R., Treyer, S., Turley, C., 2015. Contrasting futures for ocean and society from different anthropogenic CO_2 emissions scenarios. Science 349, aac4722. https://doi.org/10.1126/science.aac4722

Gavrilets, S., Vose, A., Barluenga, M., Salzburger, W., Meyer, A., 2007. Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake. Mol. Ecol. 16, 2893–2909. https://doi.org/10.1111/j.1365-294X.2007.03305.x

Geiger, D.L., Marshall, B.A., Ponder, W.F., Sasaki, T., Warén, A., 2007. Techniques for collecting, handling, preparing, storing and examining small molluscan specimens. Molluscan Res. 27, 1–50.

Geller, J., Meyer, C., Parker, M., Hawk, H., 2013. Redesign of PCR primers for mitochondrial cytochrome *c* oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol. Ecol. Resour. 13, 851–861. https://doi.org/10.1111/1755-0998.12138

Gemmell, M.R., Trewick, S.A., Crampton, J.S., Vaux, F., Hills, S.F.K., Daly, E.E., Marshall, B.A., Beu, A.G., Morgan-Richards, M., 2018. Genetic structure and shell shape variation within a rocky shore whelk suggest both diverging and constraining selection with gene flow. Biol. J. Linn. Soc. 125, 827–843. https://doi.org/10.1093/biolinnean/bly142

Ghiselli, F., Gomes-dos-Santos, A., Adema, C.M., Lopes-Lima, M., Sharbrough, J., Boore, J.L., 2021. Molluscan mitochondrial genomes break the rules. Philos. Trans. R. Soc. B Biol. Sci. 376, 20200159. https://doi.org/10.1098/rstb.2020.0159

Gilmer, R.W., Harbison, G.R., 1986. Morphology and field behavior of pteropod molluscs: feeding methods in the families Cavoliniidae, Limacinidae and Peraclididae (Gastropoda: Thecosomata). Mar. Biol. 91, 47–57. https://doi.org/10.1007/BF00397570

Gleason, L.U., Burton, R.S., 2016. Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*. Mol. Ecol. 25, 3557–3573. https://doi.org/10.1111/mec.13703

Glenn, T.C., Faircloth, B.C., 2016. Capturing Darwin's dream. Mol. Ecol. Resour. 16, 1051–1058. https://doi.org/10.1111/1755-0998.12574

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., Leproust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. 27, 182–189. https://doi.org/10.1038/nbt.1523.Solution

Goetze, E., 2005. Global population genetic structure and biogeography of the oceanic copepods *Eucalanus hyalinus* and *E. spinifer*. Evolution 59, 2378–2398. https://doi.org/10.1111/j.0014-3820.2005.tb00948.x

Goetze, E., Andrews, K.R., Peijnenburg, K.T.C.A., Portner, E., Norton, E.L., 2015. Temporal stability of genetic structure in a mesopelagic copepod. PLoS One 10, e0136087. https://doi.org/10.1371/journal.pone.0136087

Goetze, E., Hüdepohl, P.T., Chang, C., Van Woudenberg, L., Iacchei, M., Peijnenburg, K.T.C.A., 2017. Ecological dispersal barrier across the equatorial Atlantic in a migratory planktonic copepod. Prog. Oceanogr. 158, 203–212. https://doi.org/10.1016/j.pocean.2016.07.001

Golin, A.P., Choi, D., Ghahary, A., 2020. Hand sanitizers: A review of ingredients, mechanisms of action, modes of delivery, and efficacy against coronaviruses. Am. J. Infect. Control 48, 1062–1067. https://doi.org/10.1016/j.ajic.2020.06.182

Gompel, N., Prud'homme, B., Wittkopp, P.J., Kassner, V.A., Carroll, S.B., 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature 433, 481–487.

Graves, J.E., McDowell, J.R., 2015. Population structure of istiophorid billfishes. Fish. Res. 166, 21–28. https://doi.org/10.1016/j.fishres.2014.08.016

Gregory, T.R., 2019. Animal Genome Size Database [WWW Document]. URL http://www.genomesize.com (accessed 11.28.19).

Groot, C., Margolis, L., 1991. Pacific salmon life histories. UBC Press, Vancouver.

Gruber, N., 2011. Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 369, 1980–1996. https://doi.org/10.1098/rsta.2011.0003

Gruber, N., Clement, D., Carter, B.R., Feely, R.A., van Heuven, S., Hoppema, M., Ishii, M., Key, R.M., Kozyr, A., Lauvset, S.K., Lo Monaco, C., Mathis, J.T., Murata, A., Olsen, A., Perez, F.F., Sabine, C.L., Tanhua, T., Wanninkhof, R., 2019. The oceanic sink for anthropogenic CO_2 from 1994 to 2007. Science 363, 1193–1199. https://doi.org/10.1126/science.aau5153

Guerra-Varela, J., Colson, I., Backeljau, T., Breugelmans, K., Hughes, R.N., Rolán-Alvarez, E., 2009. The evolutionary mechanism maintaining shell shape and molecular differentiation between two ecotypes of the dogwhelk *Nucella lapillus*. Evol. Ecol. 23, 261–280. https://doi.org/10.1007/s10682-007-9221-5

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29, 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Halbert, K.M.K., Goetze, E., Carlon, D.B., 2013. High cryptic diversity across the global range of the migratory planktonic copepods *Pleuromamma piseki* and *P. gracilis*. PLoS One 8, e77011. https://doi.org/10.1371/journal.pone.0077011

Hall, A.C., Sherlock, E., Sykes, D., 2014. Does Micro-CT scanning damage DNA in museum specimens? J. Nat. Sci. Collect. 2, 22–29.

Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: Paleontological statistics software package for education and data analysis. Paleontol. Electron. 4, 9.

Hanlon, R.T., Messenger, J.B., 1998. Cephalopod behaviour. Cambridge University Press, Cambridge.

Hansson, L.A., 2000. Induced pigmentation in zooplankton: A trade-off between threats from predation and ultraviolet radiation. Proc. R. Soc. B Biol. Sci. 267, 2327–2331. https://doi.org/10.1098/rspb.2000.1287

Hartin, C.A., Bond-Lamberty, B., Patel, P., Mundra, A., 2016. Ocean acidification over the next three centuries using a simple global climate carbon-cycle model: Projections and sensitivities. Biogeosciences 13, 4329–4342. https://doi.org/10.5194/bg-13-4329-2016

Harvey, B.P., Al-Janabi, B., Broszeit, S., Cioffi, R., Kumar, A., Aranguren-Gassis, M., Bailey, A., Green, L., Gsottbauer, C.M., Hall, E.F., Lechler, M., Mancuso, F.P., Pereira, C.O., Ricevuto, E., Schram, J.B., Stapp, L.S., Stenberg, S., Santa Rosa, L.T., 2014. Evolution of marine organisms under climate change at different levels of biological organisation. Water (Switzerland) 6, 3545–3574. https://doi.org/10.3390/w6113545

Hays, G.C., 2003. A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. Hydrobiologia 503, 163–170. https://doi.org/10.1023/B:HYDR.0000008476.23617.b0

Hays, G.C., Richardson, A.J., Robinson, C., 2005. Climate change and marine plankton. Trends Ecol. Evol. 20, 337–344. https://doi.org/10.1016/j.tree.2005.03.004

Hellberg, M.E., 2009. Gene flow and isolation among populations of marine animals. Annu. Rev. Ecol. Evol. Syst. 40, 291–310. https://doi.org/10.1146/annurev.ecolsys.110308.120223

Hellberg, M.E., Vacquier, V.D., 1999. Rapid evolution of fertilization selectivity and lysin cDNA sequences in Teguline gastropods. Mol. Biol. Evol. 16, 839–848. https://doi.org/10.1093/oxfordjournals.molbev.a026168

Hewitt, G.M., 2011. Quaternary phylogeography: The roots of hybrid zones. Genetica 139, 617–638. https://doi.org/10.1007/s10709-011-9547-3

Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., Pitteloud, C., Dincă, V., Alvarez, N., Vila, R., 2019. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly *Thymelicus sylvestris*. Mol. Ecol. 28, 3857–3868. https://doi.org/10.1111/mec.15153

Hirai, J., 2020. Insights into reproductive isolation within the pelagic copepod *Pleuromamma abdominalis* with high genetic diversity using genome-wide SNP data. Mar. Biol. 167, 1. https://doi.org/10.1007/s00227-019-3618-x

Hirai, J., Tsuda, A., Goetze, E., 2015. Extensive genetic diversity and endemism across the global range of the oceanic copepod *Pleuromamma abdominalis*. Prog. Oceanogr. 138, 77–90. https://doi.org/10.1016/j.pocean.2015.09.002

Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K., Knowlton, N., Eakin, C.M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R.H., Dubi, A., Hatziolos, M.E., 2007. Coral reefs under rapid climate change and ocean acidification. Science 318, 1737–1742. https://doi.org/10.1126/science.1152509

Hoeh, W.R., Blakley, K.H., Brown, W.M., 1991. Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial DNA. Science 251, 1488–1490. https://doi.org/10.1126/science.1672472

Hoffman, J.I., Peck, L.S., Hillyard, G., Zieritz, A., Clark, M.S., 2010. No evidence for genetic differentiation between Antarctic limpet *Nacella concinna* morphotypes. Mar. Biol. 157, 765–778. https://doi.org/10.1007/s00227-009-1360-5

Hofmann, G.E., Barry, J.P., Edmunds, P.J., Gates, R.D., David, A., Klinger, T., Sewell, M.A., Hofmann, G.E., Barry, J.P., Edmunds, P.J., Gates, R.D., Hutchins, D.A., Klinger, T., Sewell, M.A., 2010. The effect of ocean acidification on calcifying organisms in marine ecosystems: An organism-to-ecosystem perspective. Annu. Rev. Ecol. Evol. Syst. 41, 127–147. https://doi.org/10.1146/annurev.ecolsys.ll0308.120227

Hollander, J., Butlin, R.K., 2010. The adaptive value of phenotypic plasticity in two ecotypes of a marine gastropod. BMC Evol. Biol. 10, 333. https://doi.org/10.1186/1471-2148-10-333

Hollander, J., Collyer, M.L., Adams, D.C., Johannesson, K., 2006. Phenotypic plasticity in two marine snails: Constraints superseding life history. J. Evol. Biol. 19, 1861–1872. https://doi.org/10.1111/j.1420-9101.2006.01171.x

Holman, L.E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., Rius, M., 2021. Animals, protists and bacteria share marine biogeographic patterns. Nat. Ecol. Evol. 738–746. https://doi.org/10.1038/s41559-021-01439-7

Holsinger, K.E., Weir, B.S., 2009. Genetics in geographically structured populations: Defining, estimating and interpreting F_{ST}. Nat. Rev. Genet. 10, 639–650. https://doi.org/10.1038/nrg2611

Holt, B.G., Marx, F.G., Fritz, S.A., Lessard, J.-P., Rahbek, C., 2020. Evolutionary diversification in the marine realm: a global case study with marine mammals. Front. Biogeogr. 12, e45184. https://doi.org/10.21425/F5FBG45184

Howes, E.L., Bednaršek, N., Budenbender, J., Comeau, S., Doubleday, A., Gallager, S.M., Hopcroft, R.R., Lischka, S., Maas, A.E., Bijma, J., Gattuso, J.P., 2014. Sink and swim: A status review of thecosome pteropod culture techniques. J. Plankton Res. 36, 299–315. https://doi.org/10.1093/plankt/fbu002

Hu, Z.M., Zhang, J., Lopez-Bautista, J., Duan, D.L., 2013. Asymmetric genetic exchange in the brown seaweed *Sargassum fusiforme* (Phaeophyceae) driven by oceanic currents. Mar. Biol. 160, 1407–1414. https://doi.org/10.1007/s00227-013-2192-x

Hunt, B., Strugnell, J., Bednarsek, N., Linse, K., Nelson, R.J., Pakhomov, E., Seibel, B., Steinke, D., Würzberg, L., 2010. Poles apart: The 'bipolar' pteropod species *Limacina helicina* is genetically distinct between the Arctic and Antarctic oceans. PLoS One 5, e9835. https://doi.org/10.1371/journal.pone.0009835

Hunt, B.P.V., Pakhomov, E.A., Hosie, G.W., Siegel, V., Ward, P., Bernard, K., 2008. Pteropods in Southern Ocean ecosystems. Prog. Oceanogr. 78, 193–221. https://doi.org/10.1016/j.pocean.2008.06.001

Hutchings, P., Kupriyanova, E., 2018. Cosmopolitan polychaetes – Fact or fiction? Personal and historical perspectives. Invertebr. Syst. 32, 1–9. https://doi.org/10.1071/IS17035

Iacchei, M., Butcher, E., Portner, E., Goetze, E., 2017. It's about time: Insights into temporal genetic patterns in oceanic zooplankton from biodiversity indices. Limnol. Oceanogr. 62, 1836–1852. https://doi.org/10.1002/lno.10538

Janssen, A.W., 2007. Holoplanktonic Mollusca (Gastropoda: Pterotracheoidea, Janthinoidea, Thecosomata and Gymnosomata) from the Pliocene of Pangasinan (Luzon, Philippines). Scr. Geol. 135, 29–177.

Janssen, A.W., 2012. Late Quaternary to Recent holoplanktonic Mollusca (Gastropoda) from bottom samples of the eastern Mediterranean Sea: systematics, morphology. Boll. Malacol. 48, 1–105.

Janssen, A.W., Peijnenburg, K.T.C.A., 2017. An overview of the fossil record of Pteropoda (Mollusca, Gastropoda, Heterobranchia). Cainozoic Res. 17, 3–10.

Jennings, R.M., Bucklin, A., Ossenbrügger, H., Hopcroft, R.R., 2010. Species diversity of planktonic gastropods (Pteropoda and Heteropoda) from six ocean regions based on DNA barcode analysis. Deep Sea Res. Part II Top. Stud. Oceanogr. 57, 2199–2210. https://doi.org/10.1016/j.dsr2.2010.09.022

Ji, R., Edwards, M., MacKas, D.L., Runge, J.A., Thomas, A.C., 2010. Marine plankton phenology and life history in a changing climate: Current research and future directions. J. Plankton Res. 32, 1355–1368. https://doi.org/10.1093/plankt/fbq062

Jiang, L.-Q., Carter, B.R., Feely, R.A., Lauvset, S.K., Olsen, A., 2019. Surface ocean pH and buffer capacity: past, present and future. Sci. Rep. 9, 18624. https://doi.org/10.1038/s41598-019-55039-4

Jiang, L.-Q., Feely, R.A., Carter, B.R., Greeley, D.J., Gledhill, D.K., Arzayus, K.M., 2015. Climatological distribution of aragonite saturation state in the global oceans. Global Biogeochem. Cycles 29, 1656–1673. https://doi.org/10.1002/2015GB005198

Johannesson, K., 2003. Evolution in *Littorina*: Ecology matters. J. Sea Res. 49, 107–117. https://doi.org/10.1016/S1385-1101(02)00218-6

Johannesson, K., 2009. Inverting the null-hypothesis of speciation: A marine snail perspective. Evol. Ecol. 23, 5–16. https://doi.org/10.1007/s10682-007-9225-1

Johannesson, K., 2016. What can be learnt from a snail? Evol. Appl. 9, 153–165. https://doi.org/10.1111/eva.12277

Johannesson, K., Butlin, R.K., Panova, M., Westram, A.M., 2017. Mechanisms of adaptive divergence and speciation in *Littorina saxatilis*: Integrating knowledge from ecology and genetics with new data emerging from genomic studies, in: Oleksiak, M.F., Rajora, O.P. (Eds.), Population Genomics: Marine Organisms. Springer, Cham, pp. 277–301. https://doi.org/10.1007/13836_2017_6

Johannesson, K., Moan, A. Le, Perini, S., André, C., 2020. A Darwinian laboratory of multiple contact zones. Trends Ecol. Evol. 35, 11. https://doi.org/10.1016/j.tree.2020.07.015

Johannesson, K., Panova, M., Kemppainen, P., André, C., Rolan-Alvarez, E., Butlin, R.K., 2010. Repeated evolution of reproductive isolation in a marine snail: Unveiling mechanisms of speciation. Philos. Trans. R. Soc. B Biol. Sci. 365, 1735–1747. https://doi.org/10.1098/rstb.2009.0256

Johnsen, S., 2001. Hidden in plain sight: The ecology and physiology of organismal transparency. Biol. Bull. 201, 301–318. https://doi.org/10.2307/1543609

Jones, M.R., Good, J.M., 2016. Targeted capture in evolutionary and ecological genomics. Mol. Ecol. 25, 185–202. https://doi.org/10.1111/mec.13304

Judkins, D.C., 2014. Geographical distribution of pelagic decapod shrimp in the Atlantic Ocean. Zootaxa 3895, 301–345. https://doi.org/10.11646/zootaxa.3895.3.1

Karakas, F., Wingate, J., Blanco-Bercial, L., Maas, A.E., Murphy, D.W., 2020. Swimming and sinking behavior of warm water pelagic snails. Front. Mar. Sci. 7, 749. https://doi.org/10.3389/fmars.2020.556239

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010

Kautt, A.F., Kratochwil, C.F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J., Härer, A., Hulsey, C.D., Franchini, P., Pippel, M., Myers, E.W., Meyer, A., 2020. Contrasting signatures of genomic divergence during sympatric speciation. Nature 588, 106–111. https://doi.org/10.1038/s41586-020-2845-0

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keeling, R.F., Körtzinger, A., Gruber, N., 2010. Ocean deoxygenation in a warming world. Ann. Rev. Mar. Sci. 2, 199–229. https://doi.org/10.1146/annurev.marine.010908.163855

Kender, S., McClymont, E.L., Elmore, A.C., Emanuele, D., Leng, M.J., Elderfield, H., 2016. Mid Pleistocene foraminiferal mass extinction coupled with phytoplankton evolution. Nat. Commun. 7, 11970. https://doi.org/10.1038/ncomms11970

Kess, T., Einfeldt, A.L., Wringe, B., Lehnert, S.J., Layton, K.K.S., McBride, M.C., Robert, D., Fisher, J., Le Bris, A., den Heyer, C., Shackell, N., Ruzzante, D.E., Bentzen, P., Bradbury, I.R., 2021. A putative structural variant and environmental variation associated with genomic divergence across the Northwest Atlantic in Atlantic Halibut. ICES J. Mar. Sci. 78, 2371–2384. https://doi.org/10.1093/icesjms/fsab061

Keul, N., Peijnenburg, K.T.C.A., Andersen, N., Kitidis, V., Goetze, E., Schneider, R.R., 2017. Pteropods are excellent recorders of surface temperature and carbonate ion concentration. Sci. Rep. 7, 12645. https://doi.org/10.1038/s41598-017-11708-w

Kingan, S.B., Heaton, H., Cudini, J., Lambert, C.C., Baybayan, P., Galvin, B.D., Durbin, R., Korlach, J., Lawniczak, M.K.N., Biosciences, P., Drive, O.B., Park, M., 2018. A high-quality *de novo* genome assembly from a single mosquito using PacBio sequencing. https://doi.org/10.1101/499954

Kiørboe, T., 2007. Mate finding, mating, and population dynamics in a planktonic copepod *Oithona davisae*: There are too few males. Limnol. Oceanogr. 52, 1511–1522. https://doi.org/10.4319/lo.2007.52.4.1511

Kleypas, J., 2015. Invisible barriers to dispersal. Science 348, 1086–1087. https://doi.org/10.1126/science.aab4122

Klussmann-Kolb, A., Dinapoli, A., 2006. Systematic position of the pelagic Thecosomata and Gymnosomata within Opisthobranchia (Mollusca, Gastropoda) – Revival of the Pteropoda. J. Zool. Syst. Evol. Res. 44, 118–129. https://doi.org/10.1111/j.1439-0469.2006.00351.x

Knowlton, N., 1993. Sibling species in the sea. Annu. Rev. Ecol. Syst. 24, 189–216. https://doi.org/10.1146/annurev.es.24.110193.001201

Knowlton, N., 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420, 73–90. https://doi.org/10.1023/A:1003933603879

Knowlton, N., Weight, L.A., Solórzano, L.A., Mills, D.K., Bermingham, E., 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260, 1629–1632.

Knowlton, N., Weigt, L.A., 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. R. Soc. B Biol. Sci. 265, 2257–2263. https://doi.org/10.1098/rspb.1998.0568

Koch, E.L., Morales, H.E., Larsson, J., Westram, A.M., Faria, R., Lemmon, A.R., Lemmon, E.M., Johannesson, K., Butlin, R.K., 2021. Genetic variation for adaptive traits is associated with polymorphic inversions in *Littorina saxatilis*. Evol. Lett. 5, 196–213. https://doi.org/10.1002/evl3.227

Kocot, K.M., Aguilera, F., McDougall, C., Jackson, D.J., Degnan, B.M., 2016. Sea shell diversity and rapidly evolving secretomes: insights into the evolution of biomineralization. Front. Zool. 13, 23. https://doi.org/10.1186/s12983-016-0155-z

Kohnert, P.C., Cerwenka, A.F., Brandt, A., Schrödl, M., 2020. Pteropods from the Kuril-Kamchatka Trench and the sea of Okhotsk (Euopisthobranchia; Gastropoda). Prog. Oceanogr. 181, 102259. https://doi.org/10.1016/j.pocean.2019.102259

Kollias, S., Poortvliet, M., Smolina, I., Hoarau, G., 2015. Low cost sequencing of mitogenomes from museum samples using baits capture and Ion Torrent. Conserv. Genet. Resour. 7, 345–348. https://doi.org/10.1007/s12686-015-0433-7

Koren, S., Schatz, M.C., Walenz, B.P., Martin, J., Howard, J.T., Ganapathy, G., Wang, Z., Rasko, D.A., McCombie, W.R., Jarvis, E.D., Phillippy, A.M., 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. Nat. Biotechnol. 30, 693–700. https://doi.org/10.1038/nbt.2280

Kroeker, K.J., Kordas, R.L., Crim, R., Hendriks, I.E., Ramajo, L., Singh, G.S., Duarte, C.M., Gattuso, J.P., 2013. Impacts of ocean acidification on marine organisms: Quantifying sensitivities and interaction with warming. Glob. Chang. Biol. 19, 1884–1896. https://doi.org/10.1111/gcb.12179

Krug, P.J., Vendetti, J.E., Rodriguez, A.K., Retana, J.N., Hirano, Y.M., Trowbridge, C.D., 2013. Integrative species delimitation in photosynthetic sea slugs reveals twenty candidate species in three nominal taxa studied for drug discovery, plastid symbiosis or biological control. Mol. Phylogenet. Evol. 69, 1101–1119. https://doi.org/10.1016/j.ympev.2013.07.009

Kulagin, D.N., Lunina, A.A., Simakova, U. V, Vereshchaka, A.L., 2021. Progressing diversification and biogeography of the mesopelagic *Nematoscelis* (Crustacea: Euphausiacea) in the Atlantic. ICES J. Mar. Sci. 78, 3457–3463. https://doi.org/10.1093/icesjms/fsab028

Laakkonen, H.M., Hardman, M., Strelkov, P., Väinölä, R., 2021. Cycles of trans-Arctic dispersal and vicariance, and diversification of the amphi-boreal marine fauna. J. Evol. Biol. 34, 73–96. https://doi.org/10.1111/jeb.13674

Laakmann, S., Blanco-Bercial, L., Cornils, A., 2020. The crossover from microscopy to genes in marine diversity: from species to assemblages in marine pelagic copepods. Philos. Trans. R. Soc. B Biol. Sci. 375, 20190446. https://doi.org/10.1098/rstb.2019.0446

Lachance, J., Tishkoff, S.A., 2013. SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. BioEssays 35, 780–786. https://doi.org/10.1002/bies.201300014

Laibl, C.F., Schrödl, M., Kohnert, P.C., 2019. 3D-microanatomy of a keystone planktonic species, the northern polar pteropod *Limacina helicina helicina* (Gastropoda: Heterobranchia). J. Molluscan Stud. 85, 133–142. https://doi.org/10.1093/mollus/eyy063

Lal, M.M., Southgate, P.C., Jerry, D.R., Bosserelle, C., Zenger, K.R., 2017. Swept away: ocean currents and seascape features influence genetic structure across the 18,000 Km Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster *Pinctada margaritifera*. BMC Genomics 18, 66. https://doi.org/10.1186/s12864-016-3410-y

Lalli, C.M., Conover, R.J., 1976. Microstructure of the veliger shells of gymnosomatous pteropods (Gastropoda: Opisthobranchia). Veliger 18, 237–240.

Lalli, C.M., Gilmer, R.W., 1989. Pelagic snails: The biology of holoplanktonic gastropod molluscs. Stanford University Press, California.

Lalli, C.M., Wells, F.E., 1978. Reproduction in the genus *Limacina* (Ophistobranchia: Thecosomata). J. Zool. 186, 95–108.

Lambeck, K., Chappell, J., 2001. Sea level change through the last glacial cycle. Science 292, 679–686. https://doi.org/10.1126/science.1059549

Laso-Jadart, R., Sykulski, A.M., Ambroise, C., Madoui, A., 2021. How marine currents and environment shape plankton genomic differentiation: a mosaic view from Tara Oceans metagenomic data. bioRxiv 2021.04.29.441957.

Lefébure, T., Douady, C.J., Gouy, M., Gibert, J., 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: Proposal of a molecular threshold to help species delimitation. Mol. Phylogenet. Evol. 40, 435–447. https://doi.org/10.1016/j.ympev.2006.03.014

Leigh, E.G., O'Dea, A., Vermeij, G.J., 2014. Historical biogeography of the Isthmus of Panama. Biol. Rev. 89, 148–172. https://doi.org/10.1111/brv.12048

Leigh, J.W., Bryant, D., 2015. POPART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410

LeMay, G.A., Agnarsson, I., 2020. New species of smiley-faced spider *Spintharus* (Araneae, Theridiidae) from Brazil, and comments on unobserved diversity in South America. Zookeys 915, 17–24. https://doi.org/10.3897/zookeys.915.47563

Leray, M., Knowlton, N., 2016. Censusing marine eukaryotic diversity in the twenty-first century. Philos. Trans. R. Soc. B Biol. Sci. 371, 20150331. https://doi.org/10.1098/rstb.2015.0331

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front. Zool. 10, 34. https://doi.org/10.1186/1742-9994-10-34

Lessios, H.A., 2007. Reproductive isolation between species of sea urchins. Bull. Mar. Sci. 81, 191–208.
Levin, L.A., 2006. Recent progress in understanding larval dispersal: New directions and digressions. Integr. Comp. Biol. 46, 282–297. https://doi.org/10.1093/icb/icj024
Levin, S.A., Segel, L.A., 1976. Hypothesis for origin of planktonic patchiness. Nature 259, 659.
Li, C., Hofreiter, M., Straube, N., Corrigan, S., Naylor, G.J.P., 2013. Capturing protein-coding genes across highly divergent species. Biotechniques 54, 321–326. https://doi.org/10.2144/000114039
Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arxiv Prepr. 00, 1–3.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352
Lischka, S., Büdenbender, J., Boxhammer, T., Riebesell, U., 2011. Impact of ocean acidification and elevated temperatures on early juveniles of the polar shelled pteropod *Limacina helicina*: Mortality, shell degradation, and shell growth. Biogeosciences 8, 919–932. https://doi.org/10.5194/bg-8-919-2011
Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography 20, PA1003. https://doi.org/10.1029/2004PA001071
Litsios, G., Sims, C.A., Wüest, R.O., Pearman, P.B., Zimmermann, N.E., Salamin, N., 2012. Mutualism with sea anemones triggered the adaptive radiation of clownfishes. BMC Evol. Biol. 12, 212. https://doi.org/10.1186/1471-2148-12-212
Liu, X., Fu, Y.-X., 2020. Stairway Plot 2: demographic history inference with folded SNP frequency spectra. Genome Biol. 21, 280. https://doi.org/10.1186/s13059-020-02196-9
Lohman, D.J., De Bruyn, M., Page, T., Von Rintelen, K., Hall, R., Ng, P.K.L., Shih, H. Te, Carvalho, G.C., Von Rintelen, T., 2011. Biogeography of the Indo-Australian Archipelago. Annu. Rev. Ecol. Evol. Syst. 42, 205–228. https://doi.org/10.1146/annurev-ecolsys-102710-145001
Longhurst, A.R., 2007. Ecological geography of the sea. Academic Press, San Diego. https://doi.org/10.1016/B978-0-12-455521-1.X5000-1
Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F., Storfer, A., 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol. Ecol. Resour. 17, 142–152. https://doi.org/10.1111/1755-0998.12635
Luttikhuizen, P.C., Drent, J., Baker, A.J., 2003. Disjunct distribution of highly diverged mitochondrial lineage clade and population subdivision in a marine bivalve with pelagic larval dispersal. Mol. Ecol. 12, 2215–2229. https://doi.org/10.1046/j.1365-294X.2003.01872.x
Maas, A.E., Blanco-Bercial, L., Lawson, G.L., 2013. Reexamination of the species assignment of Diacavolinia pteropods using DNA barcoding. PLoS One 8, e53889. https://doi.org/10.1371/journal.pone.0053889
Maas, A.E., Lawson, G.L., Bergan, A.J., Tarrant, A.M., 2018. Exposure to CO_2 influences metabolism, calcification, and gene expression of the thecosome pteropod *Limacina retroversa*. J. Exp. Biol. 221, jeb.164400. https://doi.org/10.1242/jeb.164400
Maas, A.E., Wishner, K.F., Seibel, B.A., 2012. Metabolic suppression in thecosomatous pteropods as an effect of low temperature and hypoxia in the eastern tropical North Pacific. Mar. Biol. 159, 1955–1967. https://doi.org/10.1007/s00227-012-1982-x
Magnan, A.K., Pörtner, H.-O., Duvat, V.K.E., Garschagen, M., Guinder, V.A., Zommers, Z., Hoegh-Guldberg, O., Gattuso, J.-P., 2021. Estimating the global risk of anthropogenic climate change. Nat. Clim. Chang. 11, 879–885. https://doi.org/10.1038/s41558-021-01156-w
Maisano Delser, P., Corrigan, S., Hale, M., Li, C., Veuille, M., Planes, S., Naylor, G., Mona, S., 2016. Population genomics of *C. melanopterus* using target gene capture data: Demographic inferences and conservation perspectives. Sci. Rep. 6, 33753. https://doi.org/10.1038/srep33753
Malinsky, M., Trucchi, E., Lawson, D.J., Falush, D., 2018. RADpainter and fineRADstructure: Population inference from RADseq data. Mol. Biol. Evol. 35, 1284–1290. https://doi.org/10.1093/molbev/msy023
Mallet, J., 2008. Hybridization, ecological races and the nature of species: Empirical evidence for the ease of speciation. Philos. Trans. R. Soc. B Biol. Sci. 363, 2971–2986. https://doi.org/10.1098/rstb.2008.0081
Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for next-generation sequencing. Nat. Methods 7, 111–118. https://doi.org/10.1038/nmeth.1419

Mani, G.S., Clarke, B.C., 1990. Mutational order: A major stochastic process in evolution. Proc. R. Soc. B Biol. Sci. 240, 29–37. https://doi.org/10.1098/rspb.1990.0025

Mann, K., Jackson, D., 2014. Characterization of the pigmented shell-forming proteome of the common grove snail *Cepaea nemoralis*. BMC Genomics 15, 249. https://doi.org/10.1186/1471-2164-15-249

Manno, C., Bednaršek, N., Tarling, G.A., Peck, V.L., Comeau, S., Adhikari, D., Bakker, D.C.E., Bauerfeind, E., Bergan, A.J., Berning, M.I., Buitenhuis, E., Burridge, A.K., Chierici, M., Flöter, S., Fransson, A., Gardner, J., Howes, E.L., Keul, N., Kimoto, K., Kohnert, P., Lawson, G.L., Lischka, S., Maas, A., Mekkes, L., Oakes, R.L., Pebody, C., Peijnenburg, K.T.C.A., Seifert, M., Skinner, J., Thibodeau, P.S., Wall-Palmer, D., Ziveri, P., 2017. Shelled pteropods in peril: Assessing vulnerability in a high CO_2 ocean. Earth-Science Rev. 169, 132–145. https://doi.org/10.1016/j.earscirev.2017.04.005

Manno, C., Rumolo, P., Barra, M., D'Albero, S., Basilone, G., Genovese, S., Mazzola, S., Bonanno, A., 2018. Condition of pteropod shells near a volcanic CO_2 vent region. Mar. Environ. Res. 143, 39–48. https://doi.org/10.1016/j.marenvres.2018.11.003

Manno, C., Tirelli, V., Accornero, A., Fonda Umani, S., 2010. Importance of the contribution of *Limacina helicina* faecal pellets to the carbon pump in Terra Nova Bay (Antarctica). J. Plankton Res. 32, 145–152. https://doi.org/10.1093/plankt/fbp108

Marçais, G., Kingsford, C., 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27, 764–770. https://doi.org/10.1093/bioinformatics/btr011

Mariani, S., Peijnenburg, K.T.C.A., Weetman, D., 2012. Independence of neutral and adaptive divergence in a low dispersal marine mollusc. Mar. Ecol. Prog. Ser. 446, 173–187. https://doi.org/10.3354/meps09507

Marko, P.B., 2004. 'What's larvae got to do with it?' Disparate patterns of post-glacial population structure in two benthic marine gastropods with identical dispersal potential. Mol. Ecol. 13, 597–611. https://doi.org/10.1046/j.1365-294X.2004.02096.x

Marlétaz, F., Le Parco, Y., Liu, S., Peijnenburg, K.T.C.A., 2017. Extreme mitogenomic variation in natural populations of chaetognaths. Genome Biol. Evol. 9, 1374–1384. https://doi.org/10.1093/gbe/evx090

Marshall, D.J., Monro, K., Bode, M., Keough, M.J., Swearer, S., 2010. Phenotype-environment mismatches reduce connectivity in the sea. Ecol. Lett. 13, 128–140. https://doi.org/10.1111/j.1461-0248.2009.01408.x

Marshall, D.J., Morgan, S.G., 2011. Ecological and evolutionary consequences of linked life-history stages in the sea. Curr. Biol. 21, R718–R725. https://doi.org/10.1016/j.cub.2011.08.022

Martin, S.H., Davey, J.W., Salazar, C., Jiggins, C.D., 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 17, e2006288. https://doi.org/10.1371/journal.pbio.2006288

Mastrantonio, V., Urbanelli, S., Porretta, D., 2019. Ancient hybridization and mtDNA introgression behind current paternal leakage and heteroplasmy in hybrid zones. Sci. Rep. 9, 19177. https://doi.org/10.1038/s41598-019-55764-w

Matz, M., 2019. Genome-wide de novo genotyping with 2bRAD [WWW Document]. URL https://github.com/z0on/2bRAD_denovo (accessed 11.28.19).

Mayr, E., 1954. Geographic speciation in tropical echnoids. Evolution 8, 1–18. https://doi.org/10.1111/j.1558-5646.1954.tb001 04.x

Mayr, E., 1963. Animal Species and Evolution. Havard University Press, Cambridge, MA. https://doi.org/10.4159/harvard.9780674865327

Mazet, O., Rodríguez, W., Grusea, S., Boitard, S., Chikhi, L., 2016. On the importance of being structured: Instantaneous coalescence rates and human evolution-lessons for ancestral population size inference? Heredity 116, 362–371. https://doi.org/10.1038/hdy.2015.104

McClary, D., Barker, M., 1998. Reproductive isolation? Interannual variability in the timing of reproduction in sympatric sea urchins, genus *Pseudechinus*. Invertebr. Biol. 117, 75–93.

McClymont, E.L., Sosdian, S.M., Rosell-Melé, A., Rosenthal, Y., 2013. Pleistocene sea-surface temperature evolution: Early cooling, delayed glacial intensification, and implications for the mid-Pleistocene climate transition. Earth-Science Rev. 123, 173–193. https://doi.org/10.1016/j.earscirev.2013.04.006

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66, 526–538. https://doi.org/10.1016/j.ympev.2011.12.007

McCormack, J.E., Tsai, W.L.E., Faircloth, B.C., 2016. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16, 1189–1203. https://doi.org/10.1111/1755-0998.12466

McDougall, C., Degnan, B.M., 2018. The evolution of mollusc shells. Wiley Interdiscip. Rev. Dev. Biol. 7, e313. https://doi.org/10.1002/wdev.313

McGinty, N., Barton, A.D., Finkel, Z. V, Johns, D.G., Irwin, A.J., 2021. Niche conservation in copepods between ocean basins. Ecography 44, 1653–1664. https://doi.org/10.1111/ecog.05690

McGowan, J.A., 1963. Geographical variation in *Limacina helicina* in the North Pacific. Syst. Assoc. Publ. 5, 109–128.

McGowan, J.A., 1967. Distributional atlas of pelagic molluscs in the California Current region. CalCOFI Atlas 6.

McGowan, J.A., 1968. The Thecosomata and Gymnosomata of California. Veliger 3, 103–129.

McKenna, A., Hanna, M., Banks, E., Mckenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

McLaren, I.A., 1963. Effects of temperature on growth of zooplankton, and the adaptive value of vertical migration. J. Fish. Res. Board Canada 20, 685–727. https://doi.org/10.1139/f63-046

McManus, G.B., Katz, L.A., 2009. Molecular and morphological methods for identifying plankton: what makes a successful marriage? J. Plankton Res. 31, 1119–1129. https://doi.org/10.1093/plankt/fbp061

Meisenheimer, J., 1905. Die Pteropoden der Deutsche Südpolar Expedition 1901-1903.

Mekkes, L., Renema, W., Bednaršek, N., Alin, S.R., Feely, R.A., Huisman, J., Roessingh, P., Peijnenburg, K.T.C.A., 2021a. Pteropods make thinner shells in the upwelling region of the California Current Ecosystem. Sci. Rep. 11, 1731. https://doi.org/10.1038/s41598-021-81131-9

Mekkes, L., Sepúlveda-Rodríguez, G., Bielkinaitė, G., Wall-Palmer, D., Brummer, G.-J.A., Dämmer, L.K., Huisman, J., van Loon, E., Renema, W., Peijnenburg, K.T.C.A., 2021b. Effects of ocean acidification on calcification of the sub-Antarctic pteropod *Limacina retroversa*. Front. Mar. Sci. 8, 581432. https://doi.org/10.3389/fmars.2021.581432

Miglietta, M.P., Faucci, A., Santini, F., 2011. Speciation in the sea: Overview of the symposium and discussion of future directions. Integr. Comp. Biol. 51, 449–455. https://doi.org/10.1093/icb/icr024

Milá, B., Van Tassell, J.L., Calderón, J.A., Rüber, L., Zardoya, R., 2017. Cryptic lineage divergence in marine environments: genetic differentiation at multiple spatial and temporal scales in the widespread intertidal goby *Gobiosoma bosc*. Ecol. Evol. 7, 5514–5523. https://doi.org/10.1002/ece3.3161

Miller, A.D., Hoffmann, A.A., Tan, M.H., Young, M., Ahrens, C., Cocomazzo, M., Rattray, A., Ierodiaconou, D.A., Treml, E., Sherman, C.D.H., 2019. Local and regional scale habitat heterogeneity contribute to genetic adaptation in a commercially important marine mollusc (*Haliotis rubra*) from southeastern Australia. Mol. Ecol. 28, 3053–3072. https://doi.org/10.1111/mec.15128

Miller, D.D., Ota, Y., Sumaila, U.R., Cisneros-Montemayor, A.M., Cheung, W.W.L., 2018. Adaptation strategies to climate change in marine systems. Glob. Chang. Biol. 24, e1–e14. https://doi.org/10.1111/gcb.13829

Miya, M., Mishida, M., 1997. Speciation in the open ocean. Nature 389, 803–804. https://doi.org/10.1038/39774

Miyamoto, H., Machida, R.J., Nishida, S., 2012. Global phylogeography of the deep-sea pelagic chaetognath *Eukrohnia hamata*. Prog. Oceanogr. 104, 99–109. https://doi.org/10.1016/j.pocean.2012.06.003

Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.-B., Norkko, A., Merilä, J., 2017. Extraordinarily rapid speciation in a marine fish. Proc. Natl. Acad. Sci. 114, 6074–6079. https://doi.org/10.1073/pnas.1615109114

Morales, H.E., Faria, R., Johannesson, K., Larsson, T., Panova, M., Westram, A.M., Butlin, R.K., 2018. Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. bioRxiv. https://doi.org/10.1101/447854

Morard, R., Lejzerowicz, F., Darling, K.F., Lecroq-Bennet, B., Pedersen, M.W., Orlando, L., Pawlowski, J., Mulitza, S., De Vargas, C., Kucera, M., 2017. Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroecology. Biogeosciences 14, 2741–2754. https://doi.org/10.5194/bg-14-2741-2017

Morton, J.E., 1954. The pelagic Mollusca of Benguela current Part I. First survey, R.R.S 'William Scoresby', March 1950 With an account of the reproductive system and sexual succession of *Limacina bulimoides*, in: Discovery Reports. University Press, Cambridge, pp. 163–200.

Moya, A., Howes, E.L., Lacoue-Labarthe, T., Forêt, S., Hanna, B., Medina, M., Munday, P.L., Ong, J.S., Teyssié, J.L., Torda, G., Watson, S.A., Miller, D.J., Bijma, J., Gattuso, J.P., 2016. Near-future pH conditions severely impact calcification, metabolism and the nervous system in the pteropod *Heliconoides inflatus*. Glob. Chang. Biol. 22, 3888–3900. https://doi.org/10.1111/gcb.13350

Mucci, A., 1983. The solubility of calcite and aragonite in seawater at various salinities, temperatures and one atmosphere total pressure. Am. J. Sci. 283, 780–799. https://doi.org/10.2475/ajs.283.7.780

Munday, P.L., Warner, R.R., Monro, K., Pandolfi, J.M., Marshall, D.J., 2013. Predicting evolutionary responses to climate change in the sea. Ecol. Lett. 16, 1488–1500. https://doi.org/10.1111/ele.12185

Murphy, D.W., Adhikari, D., Webster, D.R., Yen, J., 2016. Underwater flight by the planktonic sea butterfly. J. Exp. Biol. 219, 535–543. https://doi.org/10.1242/jeb.129205

Ng, T.P.T., Saltin, S.H., Davies, M.S., Johannesson, K., Stafford, R., Williams, G.A., 2013. Snails and their trails: The multiple functions of trail-following in gastropods. Biol. Rev. 88, 683–700. https://doi.org/10.1111/brv.12023

Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F., Bekkevold, D., 2009. Population genomics of marine fishes: Identifying adaptive variation in space and time. Mol. Ecol. 18, 3128–3150. https://doi.org/10.1111/j.1365-294X.2009.04272.x

Niemi, A., Bednaršek, N., Michel, C., Feely, R.A., Williams, W., Azetsu-Scott, K., Walkusz, W., Reist, J.D., 2021. Biological impact of ocean acidification in the Canadian Arctic: Widespread severe pteropod shell dissolution in Amundsen Gulf. Front. Mar. Sci. 8, 600184. https://doi.org/10.3389/fmars.2021.600184

Noji, T.T., Bathmann, U. V., Von Bodungen, B., Voss, M., Antia, A., Krumbholz, M., Klein, B., Peeken, I., Noji, C.I.M., Rey, F., 1997. Clearance of picoplankton-sized particles and formation of rapidly sinking aggregates by the pteropod, *Limacina retroversa*. J. Plankton Res. 19, 863–875. https://doi.org/10.1093/plankt/19.7.863

Norris, R.D., 2000. Pelagic species diversity, biogeography, and evolution. Paleobiology 26, 236–258. https://doi.org/doi:10.1017/s0094837300026956

Norris, R.D., Hull, P.M., 2012. The temporal dimension of marine speciation. Evol. Ecol. 26, 393–415. https://doi.org/10.1007/s10682-011-9488-4

Norton, E.L., Goetze, E., 2013. Equatorial dispersal barriers and limited population connectivity among oceans in a planktonic copepod. Limnol. Oceanogr. 58, 1581–1596. https://doi.org/10.4319/lo.2013.58.5.1581

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., De Queiroz, A., Farris, D.W., Norris, R.D., Stallard, R.F., Woodburne, M.O., Aguilera, O., Aubry, M.P., Berggren, W.A., Budd, A.F., Cozzuol, M.A., Coppard, S.E., Duque-Caro, H., Finnegan, S., Gasparini, G.M., Grossman, E.L., Johnson, K.G., Keigwin, L.D., Knowlton, N., Leigh, E.G., Leonard-Pingel, J.S., Marko, P.B., Pyenson, N.D., Rachello-Dolmen, P.G., Soibelzon, E., Soibelzon, L., Todd, J.A., Vermeij, G.J., Jackson, J.B.C., 2016. Formation of the Isthmus of Panama. Sci. Adv. 2, e1600883. https://doi.org/10.1126/sciadv.1600883

Obura, D.O., 2016. An Indian Ocean centre of origin revisited: Palaeogene and Neogene influences defining a biogeographic realm. J. Biogeogr. 43, 229–242. https://doi.org/10.1111/jbi.12656

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2019. vegan: Community Ecology Package.

Oliveira-Koblitz, V.S. de, Larrazábal, M.E. de L., 2014. Characterization of the geographical distribution pattern of the family Limacinidae Gray, 1840 (Mollusca – Gastropoda) in the waters of Northeastern of Brazil. Biota Neotrop. 14, e20130029. https://doi.org/10.1590/1676-06032014002913

Omori, M., Hamner, W.M., 1982. Patchy distribution of zooplankton: Behavior, population assessment and sampling problems. Mar. Biol. 72, 193–200. https://doi.org/10.1007/BF00396920

Orr, H.A., Unckless, R.L., 2008. Population extinction and the genetics of adaptation. Am. Nat. 172, 160–169. https://doi.org/10.1086/589460

Orr, J.C., Fabry, V.J., Aumont, O., Bopp, L., Doney, S.C., Feely, R.M., Gnanadesikan, A., Gruber, N., Ishida, A., Key, R.M., Lindsay, K., Maier-reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R.G., Plattner, G.K., Rodgers, K.B., Sabine, C.L., Sarmiento, J.L., Schlitzer, R., Slater, R.D., Totterdell, I.J., Weirig, M.F., Yamanaka, Y., Yool, A., Matear, Richard, 2005. Anthropogenic decline in high-latitude ocean carbonate by 2100. Nature 437, 681–686. https://doi.org/10.1038/nature04095.

Ørsted, M., Hoffmann, A.A., Sverrisdóttir, E., Nielsen, K.L., Kristensen, T.N., 2019. Genomic variation predicts adaptive evolutionary responses better than population bottleneck history. PLoS Genet. 15, e1008205. https://doi.org/10.1371/journal.pgen.1008205

Padgham, M., Sumner, M.D., 2020. geodist: Fast, Dependency-Free Geodesic Distance Calculations.

Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. Front. Zool. 7, 16. https://doi.org/10.1186/1742-9994-7-16

Paijmans, J.L.A., Fickel, J., Courtiol, A., Hofreiter, M., Förster, D.W., 2016. Impact of enrichment conditions on cross-species capture of fresh and degraded DNA. Mol. Ecol. Resour. 16, 42–55. https://doi.org/10.1111/1755-0998.12420

Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. Annu. Rev. Ecol. Syst. 25, 547–572.

Palumbi, S.R., 2004. Marine reserves and Ocean neighborhoods: The spatial scale of marine populations and their management. Annu. Rev. Environ. Resour. 29, 31–68. https://doi.org/10.1146/annurev.energy.29.062403.102254

Pastor, M. V., Palter, J.B., Pelegrí, J.L., Dunne, J.P., 2013. Physical drivers of interannual chlorophyll variability in the eastern subtropical North Atlantic. J. Geophys. Res. Ocean. 118, 3871–3886. https://doi.org/10.1002/jgrc.20254

Peijnenburg, K.T.C.A., Breeuwer, J.A.J., Pierrot-Bults, A.C., Menken, S.B.J., 2004. Phylogeography of the planktonic chaetognath *Sagitta setosa* reveals isolation in European seas. Evolution 58, 1472–1487. https://doi.org/10.1111/j.0014-3820.2004.tb01728.x

Peijnenburg, K.T.C.A., Fauvelot, C., Breeuwer, J.A.J., Menken, S.B.J., 2006. Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol. Ecol. 15, 3319–3338. https://doi.org/10.1111/j.1365-294X.2006.03002.x

Peijnenburg, K.T.C.A., Goetze, E., 2013. High evolutionary potential of marine zooplankton. Ecol. Evol. 3, 2765–2783. https://doi.org/10.1002/ece3.644

Peijnenburg, K.T.C.A., Janssen, A.W., Wall-Palmer, D., Goetze, E., Maas, A.E., Todd, J.A., Marlétaz, F., 2020. The origin and diversification of pteropods precede past perturbations in the Earth's carbon cycle. Proc. Natl. Acad. Sci. 117, 25609–25617. https://doi.org/10.1073/pnas.1920918117

Pelejero, C., Calvo, E., Hoegh-Guldberg, O., 2010. Paleo-perspectives on ocean acidification. Trends Ecol. Evol. 25, 332–344. https://doi.org/10.1016/j.tree.2010.02.002

Pentinsaari, M., Salmela, H., Mutanen, M., Roslin, T., 2016. Molecular evolution of a widely-adopted taxonomic marker (COI) across the animal tree of life. Sci. Rep. 6, 35275. https://doi.org/10.1038/srep35275

Pérez-Portela, R., Arranz, V., Rius, M., Turon, X., 2013. Cryptic speciation or global spread? The case of a cosmopolitan marine invertebrate with limited dispersal capabilities. Sci. Rep. 3, 3197. https://doi.org/10.1038/srep03197

Pérez, V., Fernández, E., Marañón, E., Serret, P., García-Soto, C., 2005. Seasonal and interannual variability of chlorophyll *a* and primary production in the Equatorial Atlantic: *in situ* and remote sensing observations. J. Plankton Res. 27, 189–197. https://doi.org/10.1093/plankt/fbh159

Petrick, B., Martínez-García, A., Auer, G., Reuning, L., Auderset, A., Deik, H., Takayanagi, H., De Vleeschouwer, D., Iryu, Y., Haug, G.H., 2019. Glacial Indonesian Throughflow weakening across the Mid-Pleistocene Climatic Transition. Sci. Rep. 9, 16995. https://doi.org/10.1038/s41598-019-53382-0

Pfaller, J.B., Payton, A.C., Bjorndal, K.A., Bolten, A.B., McDaniel, S.F., 2019. Hitchhiking the high seas: Global genomics of rafting crabs. Ecol. Evol. 9, 957–974. https://doi.org/10.1002/ece3.4694

Philander, S.G., 2001. Atlantic Ocean Equatorial Currents. Encycl. Ocean Sci. 188–191. https://doi.org/10.1006/rwos.2001.0361

Pierrot-Bults, A.C., Van der Spoel, S., 1979. Speciation in macrozooplankton, in: Zoogeography and Diversity in Plankton. Halsted, New York, pp. 144–167.

Pineda, J., Reyns, N.B., Starczak, V.R., 2009. Complexity and simplification in understanding recruitment in benthic populations. Popul. Ecol. 51, 17–32. https://doi.org/10.1007/s10144-008-0118-0

Pinsky, M.L., Selden, R.L., Kitchel, Z.J., 2020. Climate-driven shifts in marine species ranges: Scaling from organisms to communities. Ann. Rev. Mar. Sci. 12, 153–179. https://doi.org/10.1146/annurev-marine-010419-010916

Poloczanska, E.S., Burrows, M.T., Brown, C.J., García Molinos, J., Halpern, B.S., Hoegh-Guldberg, O., Kappel, C. V., Moore, P.J., Richardson, A.J., Schoeman, D.S., Sydeman, W.J., 2016. Responses of marine organisms to climate change across oceans. Front. Mar. Sci. 3, 62. https://doi.org/10.3389/fmars.2016.00062

Portik, D.M., Smith, L.L., Bi, K., 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). Mol. Ecol. Resour. 16, 1069–1083. https://doi.org/10.1111/1755-0998.12541

Postel, U., Glemser, B., Salazar Alekseyeva, K., Eggers, S.L., Groth, M., Glöckner, G., John, U., Mock, T., Klemm, K., Valentin, K., Beszteri, B., 2020. Adaptive divergence across Southern Ocean gradients in the pelagic diatom *Fragilariopsis kerguelensis*. Mol. Ecol. 29, 4913–4924. https://doi.org/10.1111/mec.15554

Potkamp, G., Fransen, C.H.J.M., 2019. Speciation with gene flow in marine systems. Contrib. to Zool. 88, 133–172. https://doi.org/10.1163/18759866-20191344

Prada, C., Hellberg, M.E., 2014. Strong natural selection on juveniles maintains a narrow adult hybrid zone in a broadcast spawner. Am. Nat. 184, 702–713. https://doi.org/10.1086/678403

Prairie, J.C., Sutherland, K.R., Nickols, K.J., Kaltenberg, A.M., 2012. Biophysical interactions in the plankton: A cross-scale review. Limnol. Oceanogr. Fluids Environ. 2, 121–145. https://doi.org/10.1215/21573689-1964713

Prazeres, M., Morard, R., Roberts, T.E., Doo, S.S., Jompa, J., Schmidt, C., Stuhr, M., Renema, W., Kucera, M., 2020. High dispersal capacity and biogeographic breaks shape the genetic diversity of a globally distributed reef-dwelling calcifier. Ecol. Evol. 10, 5976–5989. https://doi.org/10.1002/ece3.6335

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A., Savidge, G., 2009. High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus*. Proc. R. Soc. B Biol. Sci. 276, 301–307. https://doi.org/10.1098/rspb.2008.1062

Quattrini, A.M., Faircloth, B.C., Dueñas, L.F., Bridge, T.C.L., Brugler, M.R., Calixto-Botía, I.F., DeLeo, D.M., Forêt, S., Herrera, S., Lee, S.M.Y., Miller, D.J., Prada, C., Rádis-Baptista, G., Ramírez-Portilla, C., Sánchez, J.A., Rodríguez, E., McFadden, C.S., 2018. Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. Mol. Ecol. Resour. 18, 281–295. https://doi.org/10.1111/1755-0998.12736

R Core Team, 2017. R: A language and environment for statistical computing.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. https://doi.org/10.1093/sysbio/syy032

Ramos-Silva, P., Marin, F., 2016. Proteins as functional units of biocalcification – an overview. Key Eng. Mater. 672, 183–190. https://doi.org/10.4028/www.scientific.net/KEM.672.183

Rampal, J., 2017. Euthecosomata (Mollusca, Gastropoda, Thecosomata). Taxonomic review. bioRxiv 1–82. https://doi.org/10.1101/098475

Ratnasingham, S., Hebert, P.D.N., 2007. BOLD: The Barcode Of Life Data system: Barcoding. Mol. Ecol. Notes 7, 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Ratnasingham, S., Hebert, P.D.N., 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. PLoS One 8, e66213. https://doi.org/10.1371/journal.pone.0066213

Reid, D.G., Lal, K., Mackenzie-Dodds, J., Kaligis, F., Littlewood, D.T.J., Williams, S.T., 2006. Comparative phylogeography and species boundaries in *Echinolittorina* snails in the central Indo-West Pacific. J. Biogeogr. 33, 990–1006. https://doi.org/10.1111/j.1365-2699.2006.01469.x

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., Maury, O., 2013. Dynamic biogeochemical provinces in the global ocean. Global Biogeochem. Cycles 27, 1046–1058. https://doi.org/10.1002/gbc.20089

Reygondeau, G., Maury, O., Beaugrand, G., Fromentin, J.M., Fonteneau, A., Cury, P., 2012. Biogeography of tuna and billfish communities. J. Biogeogr. 39, 114–129. https://doi.org/10.1111/j.1365-2699.2011.02582.x

Rice, E.S., Green, R.E., 2018. New Approaches for Genome Assembly and Scaffolding. Annu. Rev. Anim. Biosci. 7, 17–40. https://doi.org/10.1146/annurev-animal-020518-115344

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution 43, 223–225. https://doi.org/10.1111/j.1558-5646.1989.tb04220.x

Richter, D.J., Watteaux, R., Vannier, T., Leconte, J., Frémont, P., 2019. Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. bioRxiv 1–36. https://doi.org/10.1101/867739

Riebesell, U., Gattuso, J.P., 2015. Lessons learned from ocean acidification research. Nat. Clim. Chang. 5, 12–14. https://doi.org/10.1038/nclimate2456

Riebesell, U., Zondervan, I., Rost, B., Tortell, P.D., Zeebe, R.E., Morel, F.M.M., 2000. Reduced calcification of marine plankton in response to increased atmospheric CO_2. Nature 407, 364–366. https://doi.org/10.1038/35030078

Riginos, C., Crandall, E.D., Liggins, L., Bongaerts, P., Treml, E.A., 2016. Navigating the currents of seascape genomics: How spatial analyses can augment population genomic studies. Curr. Zool. 62, 581–601. https://doi.org/10.1093/cz/zow067

Ritcher, G., 1979. Die thecosomen Pteropoden der 'Meteor'-Expedition in den Indischen Ozean 1964/65. Forsch. Ergenbnisse 29, 1–29.

Rohlf, F.J., 2015. The tps series of software. Hystrix 26, 9–12. https://doi.org/10.4404/hystrix-26.1-11264

Rohlf, F.J., Slice, D., 1990. Extension of the procrustes method for the optimal superimposition of landmarks. Syst. Zool 39, 40–59.

Rolán, E., Guerra-Varela, J., Colson, I., Hughes, R.N., Rolán-Alvarez, E., 2004. Morphological and genetic analysis of two sympatric morphs of the dogwhelk *Nucella lapillus* (Gastropoda: Muricidae) from Galicia (Northwestern Spain). J. Molluscan Stud. 70, 179–185. https://doi.org/10.1093/mollus/70.2.179

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., Legendre, L., 2009. Global latitudinal variations in marine copepod diversity and environmental factors. Proc. R. Soc. B Biol. Sci. 276, 3053–3062. https://doi.org/10.1098/rspb.2009.0742

Roth, V.L., Mercer, J.M., 2000. Morphometrics in development and evolution. Am. Zool. 40, 801–810. https://doi.org/10.1093/icb/40.5.801

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54, 952–961. https://doi.org/10.1080/10635150500234674

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. Mol. Ecol. 24, 6241–6255. https://doi.org/10.1111/mec.13471

Sandoval-Castillo, J., Robinson, N.A., Hart, A.M., Strain, L.W.S., Beheregaray, L.B., 2018. Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. Mol. Ecol. 27, 1603–1620. https://doi.org/10.1111/mec.14526

Sanford, E., Kelly, M.W., 2011. Local adaptation in marine invertebrates. Ann. Rev. Mar. Sci. 3, 509–535. https://doi.org/10.1146/annurev-marine-120709-142756

Sayers, E.W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K.D., Karsch-Mizrachi, I., 2019. GenBank. Nucleic Acids Res. 47, D94–D99. https://doi.org/10.1093/nar/gky989

Schell, T., Feldmeyer, B., Schmidt, H., Greshake, B., Tills, O., Truebano, M., Rundle, S.D., Paule, J., Ebersberger, I., Pfenninger, M., 2017. An annotated draft genome for *Radix auricularia* (Gastropoda, Mollusca). Genome Biol. Evol. 9, 585–592. https://doi.org/10.1093/gbe/evx032

Schluter, D., 2009. Evidence for ecological speciation and its alternative. Science 323, 737–741. https://doi.org/10.1126/science.1160006

Schluter, D., Conte, G.L., 2009. Genetics and ecological speciation. Light Evol. 3, 47–64. https://doi.org/10.17226/12692

Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., Von Der Heyden, S., Toonen, R.J., 2016. A decade of seascape genetics: Contributions to basic and applied marine connectivity. Mar. Ecol. Prog. Ser. 554, 1–19. https://doi.org/10.3354/meps11792

Sgrò, C.M., Lowe, A.J., Hoffmann, A.A., 2011. Building evolutionary resilience for conserving biodiversity under climate change. Evol. Appl. 4, 326–337. https://doi.org/10.1111/j.1752-4571.2010.00157.x

Shakun, J., Carlson, A.E., 2010. A global perspective on Last Glaicial Maximum to Holocene climate change. Quat. Sci. Rev. 29, 1801–1816. https://doi.org/10.1016/j.quascirev.2010.03.016

Shaw, K.L., Mullen, S.P., 2014. Speciation continuum. J. Hered. 105, 741–742. https://doi.org/10.1093/jhered/esu060

Shimizu, K., Kimoto, K., Noshita, K., Wakita, M., Fujiki, T., Sasaki, T., 2018. Phylogeography of the pelagic snail *Limacina helicina* (Gastropoda: Thecosomata) in the subarctic western North Pacific. J. Molluscan Stud. 84, 30–37. https://doi.org/10.1093/mollus/eyx040

Shimizu, K., Noshita, K., Kimoto, K., Sasaki, T., 2021. Phylogeography and shell morphology of the pelagic snail *Limacina helicina* in the Okhotsk Sea and western North Pacific. Mar. Biodivers. 51, 22. https://doi.org/10.1007/s12526-021-01166-z

Shkoldina, L.S., 1999. On the systematics of the pteropod mollusk *Limacina helicina* from the sea of Okhotsk. Russ. J. Mar. Biol. 25, 330–336.

Simakov, O., Marletaz, F., Cho, S., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D., Larsson, T., Lv, J., Arendt, D., Savage, R., Osoegawa, K., de Jong, P., Grimwood, J., Chapman, J.A., Shapiro, H., Aerts, A., Otillar, R.P., Terry, A.Y., Boore, J.L., Grigoriev, I. V, Lindberg, D.R., Seaver, E.C., Weisblat, D.A., Putnam, N.H., Rokhsar, D.S., 2013. Insights into bilaterian evolution from three spiralian genomes. Nature 493, 526–531. https://doi.org/10.1038/nature11696

Simal Busto, D., 2019. Biogeography of *Limacina* euthecosomes: An updated review of records from 1898 – 2018 (MSc Thesis).

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J.J., Bierne, N., 2021. How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. J. Evol. Biol. 34, 208–223. https://doi.org/10.1111/jeb.13709

Singhal, S., Moritz, C., 2012. Strong selection against hybrids maintains a narrow contact zone between morphologically cryptic lineages in a rainforest lizard. Evolution 66, 1474–1489. https://doi.org/10.1111/j.1558-5646.2011.01539.x

Sloan, D.B., Havird, J.C., Sharbrough, J., 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. Mol. Ecol. 26, 2212–2236. https://doi.org/10.1111/mec.13959

Sommeria-Klein, G., Watteaux, R., Iudicone, D., Bowler, C., Morlon, H., 2020. Global drivers of eukaryotic plankton biogeography in the sunlit ocean. bioRxiv 1–25. https://doi.org/10.1101/2020.09.08.287524

Song, H., Buhay, J.E., Whiting, M.F., Crandall, K.A., 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. Proc. Natl. Acad. Sci. U. S. A. 105, 13486–13491. https://doi.org/10.1073/pnas.0803076105

Sromek, L., Lasota, R., Wolowicz, M., 2015. Impact of glaciations on genetic diversity of pelagic mollusks: Antarctic *Limacina antarctica* and Arctic *Limacina helicina*. Mar. Ecol. Prog. Ser. 525, 143–152. https://doi.org/10.3354/meps11237

Stange, M., Sánchez-Villagra, M.R., Salzburger, W., Matschiner, M., 2018. Bayesian divergence-time estimation with genome-wide single-nucleotide polymorphism data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus. Syst. Biol. 67, 681–699. https://doi.org/10.1093/sysbio/syy006

Stanke, M., Morgenstern, B., 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 33, 465–467. https://doi.org/10.1093/nar/gki458

Stankowski, S., Ravinet, M., 2021. Defining the speciation continuum. Evolution 75, 1256–1273. https://doi.org/10.1111/evo.14215

Stankowski, S., Westram, A.M., Zagrodzka, Z.B., Eyres, I., Broquet, T., Johannesson, K., Butlin, R.K., 2020. The evolution of strong reproductive isolation between sympatric intertidal snails. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 375, 20190545. https://doi.org/10.1098/rstb.2019.0545

Stanley, R.R.E., DiBacco, C., Lowen, B., Beiko, R.G., Jeffery, N.W., Van Wyngaarden, M., Bentzen, P., Brickman, D., Benestan, L., Bernatchez, L., Johnson, C., Snelgrove, P.V.R., Wang, Z., Wringe, B.F., Bradbury, I.R., 2018. A climate-associated multispecies cryptic cline in the northwest Atlantic. Sci. Adv. 4, eaaq0929. https://doi.org/10.1126/sciadv.aaq0929

Stein, E.D., White, B.P., Mazor, R.D., Miller, P.E., Pilgrim, E.M., 2013. Evaluating ethanol-based sample preservation to facilitate use of DNA barcoding in routine freshwater biomonitoring programs using benthic macroinvertebrates. PLoS One 8, e51273. https://doi.org/10.1371/journal.pone.0051273

Steinberg, D.K., Landry, M.R., 2017. Zooplankton and the ocean carbon cycle. Ann. Rev. Mar. Sci. 9, 413–444. https://doi.org/10.1146/annurev-marine-010814-015924

Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162–1169. https://doi.org/10.1086/379378

Stoeckle, M.Y., Thaler, D.S., 2014. DNA barcoding works in practice but not in (neutral) theory. PLoS One 9, e100755. https://doi.org/10.1371/journal.pone.0100755

Sulpis, O., Jeansson, E., Dinauer, A., Lauvset, S.K., Middelburg, J.J., 2021. Calcium carbonate dissolution patterns in the ocean. Nat. Geosci. 14, 423–428. https://doi.org/10.1038/s41561-021-00743-y

Sun, W., Jayaraman, S., Chen, W., Persson, K.A., Ceder, G., 2015. Nucleation of metastable aragonite $CaCO_3$ in seawater. Proc. Natl. Acad. Sci. U. S. A. 112, 3199–3204. https://doi.org/10.1073/pnas.1506100112

Sunday, J.M., Calosi, P., Dupont, S., Munday, P.L., Stillman, J.H., Reusch, T.B.H., 2014. Evolution in an acidifying ocean. Trends Ecol. Evol. 29, 117–125. https://doi.org/10.1016/j.tree.2013.11.001

Suren, H., Hodgins, K., Yeaman, S., NurkowskiKA, Smets, P., Rieseberg, L., Aitken, S., Holliday, J., 2016. Exome capture from the spruce and pine giga-genomes. Mol. Ecol. Resour. 16, 1136–1146. https://doi.org/10.1111/1755-0998.12570

Sutton, T.T., Clark, M.R., Dunn, D.C., Halpin, P.N., Rogers, A.D., Guinotte, J., Bograd, S.J., Angel, M. V., Perez, J.A.A., Wishner, K., Haedrich, R.L., Lindsay, D.J., Drazen, J.C., Vereshchaka, A., Piatkowski, U., Morato, T., Błachowiak-Samołyk, K., Robison, B.H., Gjerde, K.M., Pierrot-Bults, A., Bernal, P., Reygondeau, G., Heino, M., 2017. A global biogeographic classification of the mesopelagic zone. Deep Sea Res. Part I Oceanogr. Res. Pap. 126, 85–102. https://doi.org/10.1016/j.dsr.2017.05.006

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Takeuchi, T., 2017. Molluscan Genomics: Implications for Biology and Aquaculture. Curr. Mol. Biol. Reports 3, 297–305. https://doi.org/10.1007/s40610-017-0077-3

Tange, O., 2011. GNU Parallel: The command-line power tool.;login USENIX Mag. https://doi.org/http://dx.doi.org/10.5281/zenodo.16303

Teasdale, L.C., Köhler, F., Murray, K.D., O'Hara, T., Moussalli, A., 2016. Identification and qualification of 500 nuclear, single-copy, orthologous genes for the Eupulmonata (Gastropoda) using transcriptome sequencing and exon capture. Mol. Ecol. Resour. 16, 1107–1123. https://doi.org/10.1111/1755-0998.12552

Telesca, L., Michalek, K., Sanders, T., Peck, L.S., Thyrring, J., Harper, E.M., 2018. Blue mussel shell shape plasticity and natural environments: a quantitative approach. Sci. Rep. 8, 2865. https://doi.org/10.1038/s41598-018-20122-9

Terhaar, J., Kwiatkowski, L., Bopp, L., 2020. Emergent constraint on Arctic Ocean acidification in the twenty-first century. Nature 582, 379–383. https://doi.org/10.1038/s41586-020-2360-3

Tesch, J.J., 1946. The thecosomatous pteropods. I. The Atlantic. Copenhagen.

Tesch, J.J., 1948. The thecosomatous pteropods. II. The Indo-Pacific. Copenhagen.

Teske, P.R., Golla, T.R., Sandoval-Castillo, J., Emami-Khoyi, A., Van Der Lingen, C.D., Von Der Heyden, S., Chiazzari, B., Jansen Van Vuuren, B., Beheregaray, L.B., 2018. Mitochondrial DNA is unsuitable to test for isolation by distance. Sci. Rep. 8, 1–9. https://doi.org/10.1038/s41598-018-25138-9

Therkildsen, N.O., Wilder, A.P., Conover, D.O., Munch, S.B., Baumann, H., Palumbi, S.R., 2019. Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. Science 365, 487–490. https://doi.org/10.1126/science.aaw7271

Thibodeau, P.S., Steinberg, D.K., Maas, A.E., 2020. Effects of temperature and food concentration on pteropod metabolism along the Western Antarctic Peninsula. J. Exp. Mar. Bio. Ecol. 530–531, 151412. https://doi.org/10.1016/j.jembe.2020.151412

Thomas, J.A., Welch, J.J., Woolfit, M., Bromham, L., 2006. There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. Proc. Natl. Acad. Sci. U. S. A. 103, 7366–7371. https://doi.org/10.1073/pnas.0510251103

Thomaz, D., Guiller, A., Clarke, B., 1996. Extreme divergence of mitochondrial DNA within species of pulmonate land snails. Proc. R. Soc. B Biol. Sci. 263, 363–368. https://doi.org/10.1098/rspb.1996.0056

Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Tomasini, M., Peischl, S., 2020. When does gene flow facilitate evolutionary rescue? Evolution 74, 1640–1653. https://doi.org/10.1111/evo.14038

Treangen, T.J., Salzberg, S.L., 2012. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat. Rev. Genet. 13, 36–46. https://doi.org/10.1038/nrg3117

Trudnowska, E., Sagan, S., Kwasniewski, S., Darecki, M., Blachowiak-Samolyk, K., 2015. Fine-scale zooplankton vertical distribution in relation to hydrographic and optical characteristics of the surface waters on the Arctic shelf. J. Plankton Res. 37, 120–133. https://doi.org/10.1093/plankt/fbu087

Truelove, N.K., Kough, A.S., Behringer, D.C., Paris, C.B., Box, S.J., Preziosi, R.F., Butler, M.J., 2017. Biophysical connectivity explains population genetic structure in a highly dispersive marine species. Coral Reefs 36, 233–244. https://doi.org/10.1007/s00338-016-1516-y

Tsurumi, M., Mackas, D.L., Whitney, F.A., DiBacco, C., Galbraith, M.D., Wong, C.S., 2005. Pteropods, eddies, carbon flux, and climate variability in the Alaska Gyre. Deep. Res. Part II Top. Stud. Oceanogr. 52, 1037–1053. https://doi.org/10.1016/j.dsr2.2005.02.005

Turelli, M., Barton, N.H., Coyne, J.A., 2001. Theory of Speciation. Trends Ecol. Evol. 16, 330–343.

Turney, C.S.M., Jones, R.T., McKay, N.P., Van Sebille, E., Thomas, Z.A., Hillenbrand, C.D., Fogwill, C.J., 2020. A global mean sea surface temperature dataset for the Last Interglacial (129-116 ka) and contribution of thermal expansion to sea level change. Earth Syst. Sci. Data 12, 3341–3356. https://doi.org/10.5194/essd-12-3341-2020

Valdez-Mondragón, A., 2010. Zootaxa, Two new species of spiders of the genus *Selenops Latreille*, 1819 (Araneae: Selenopidae) and redescription of *Selenops scitus* Muma, 1953 from Mexico. Zootaxa 58, 47–58.

van der Spoel, S., 1967. Euthecosomata: a group with remarkable deveopmental stages (Gastropoda, Pteropoda) (PhD thesis).

van der Spoel, S., Bleeker, J., Kobayasi, H., 1993. From *Cavolinia longirostris* to twenty-four *Diacavolinia* taxa, with a phylogenetic discussion (Mollusca, Gastropoda). Bijdr. tot Dierkd. 62, 127–166.

van der Spoel, S., Dadon, J.R., 1999. Pteropoda, in: Boltovskoy, D. (Ed.), South Atlantic Zooplankton. Backhuys Publishers, Leiden, pp. 649–706.

van der Spoel, S., Heyman, R.P., 1983. A comparative atlas of zooplankton: Biological patterns in the oceans. Wetenschappelijke uitgeverij Bunge, Utrecht.

van der Spoel, S., Newman, L.J., Estep, K.W., 1997. Pelagic molluscs of the world. ETI World Biodivers. Database CD-ROM Ser.

van der Sprong, J., 2019. What do we know about cryptic diversity in marine zooplankton? (MSc thesis).

Vermeij, G.J., 2002. Characters in context: molluscan shells and the forces that mold them. Paleobiology 28, 41–54. https://doi.org/10.1666/0094-8373(2002)028<0041:cicmsa>2.0.co;2

Via, S., 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends Ecol. Evol. 16, 381–390.

von der Heyden, S., 2017. Making evolutionary history count: biodiversity planning for coral reef fishes and the conservation of evolutionary processes. Coral Reefs 36, 183–194. https://doi.org/10.1007/s00338-016-1512-2

Vonnemann, V., Schrödl, M., Klussmann-Kolb, A., Wägele, H., 2005. Reconstruction of the phylogeny of the Opisthobranchia (Mollusca: Gastropoda) by means of 18S and 28S rRNA gene sequences. J. Molluscan Stud. 71, 113–125. https://doi.org/10.1093/mollus/eyi014

Wall-Palmer, D., Burridge, A.K., Goetze, E., Stokvis, F.R., Janssen, A.W., Mekkes, L., Moreno-Alcántara, M., Bednaršek, N., Schiøtte, T., Sørensen, M.V., Smart, C.W., Peijnenburg, K.T.C.A., 2018. Biogeography and genetic diversity of the atlantid heteropods. Prog. Oceanogr. 160, 1–25. https://doi.org/10.1016/j.pocean.2017.11.004

Wall-Palmer, D., Burridge, A.K., Peijnenburg, K.T.C.A., Janssen, A., Goetze, E., Kirby, R., Hart, M.B., Smart, C.W., 2016a. Evidence for the validity of *Protatlanta sculpta* (Gastropoda: Pterotracheoidea). Contrib. to Zool. 85, 423–435. https://doi.org/10.1163/18759866-08504003

Wall-Palmer, D., Janssen, A., Goetze, E., Choo, L.Q., Mekkes, L., Peijnenburg, K.T.C.A., 2020. Fossil-calibrated molecular phylogeny of atlantid heteropods (Gastropoda, Pterotracheoidea). BMC Evol. Biol. 20, 124. https://doi.org/10.21203/rs.3.rs-18183/v1

Wall-Palmer, D., Smart, C.W., Hart, M.B., Leng, M.J., Conversi, A., Borghini, M., Manini, E., Aliani, S., 2014. Late Pleistocene pteropods, heteropods and planktonic foraminifera from the Caribbean Sea, Mediterranean Sea and Indian Ocean. Micropaleontology 60, 557–578.

Wall-Palmer, D., Smart, C.W., Kirby, R., Hart, M.B., Peijnenburg, K.T.C.A., Janssen, A.W., 2016b. A review of the ecology, palaeontology and distribution of atlantid heteropods (Caenogastropoda: Pterotracheoidea: Atlantidae). J. Molluscan Stud. 82, 221–234. https://doi.org/10.1093/mollus/eyv063

Waples, R.S., 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. J. Hered. 89, 438–450. https://doi.org/10.1093/jhered/89.5.438

Ward, B.A., Cael, B.B., Collins, S., Young, C.R., 2021. Selective constraints on global plankton dispersal. Proc. Natl. Acad. Sci. 118, e2007388118. https://doi.org/10.1073/pnas.2007388118

Weersing, K., Toonen, R., 2009. Population genetics, larval dispersal, and connectivity in marine systems. Mar. Ecol. Prog. Ser. 393, 1–12. https://doi.org/10.3354/meps08287

Weiner, A.K.M., Morard, R., Weinkauf, M.F.G., Darling, K.F., André, A., Quillévéré, F., Ujiie, Y., Douady, C.J., de Vargas, C., Kucera, M., 2016. Methodology for single-cell genetic analysis of planktonic foraminifera for studies of protist diversity and evolution. Front. Mar. Sci. 3, 255. https://doi.org/10.3389/fmars.2016.00255

Weiner, A.K.M., Weinkauf, M.F.G., Kurasawa, A., Darling, K.F., Kucera, M., Grimm, G.W., 2014. Phylogeography of the tropical planktonic foraminifera lineage *Globigerinella* reveals isolation inconsistent with passive dispersal by ocean currents. PLoS One 9, e92148. https://doi.org/10.1371/journal.pone.0092148

Wells, F.E., 1976. Growth rate of four species of euthecosomatous pteropods occuring off Barbados, West Indies. Nautilus 90, 114–16.

Westram, A.M., Rafajlović, M., Chaube, P., Faria, R., Larsson, T., Panova, M., Ravinet, M., Blomberg, A., Mehlig, B., Johannesson, K., Butlin, R., 2018. Clines on the seashore: The genomic architecture underlying rapid divergence in the face of gene flow. Evol. Lett. 2, 297–309. https://doi.org/10.1002/evl3.74

White, C., Selkoe, K.A., Watson, J., Siegel, D.A., Zacherl, D.C., Toonen, R.J., 2010. Ocean currents help explain population genetic structure. Proc. R. Soc. B Biol. Sci. 277, 1685–1694. https://doi.org/10.1098/rspb.2009.2214

Whitlock, M.C., Lotterhos, K.E., 2015. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. Am. Nat. 186, S24–S36. https://doi.org/10.1086/682949

Whittaker, K.A., Rynearson, T.A., 2017. Evidence for environmental and ecological selection in a microbe with no geographic limits to gene flow. Proc. Natl. Acad. Sci. U. S. A. 114, 2651–2656. https://doi.org/10.1073/pnas.1612346114

Wickham, H., 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York.

Wiernasz, D.C., Kingsolver, J.G., 1992. Wing melanin pattern mediates species recognition in *Pieris occidentalis*. Anim. Behav. 43, 89–94. https://doi.org/10.1016/S0003-3472(05)80074-0

Wilke, T., Schultheiß, R., Albrecht, C., 2009. As time goes by: A simple fool's guide to molecular clock approaches in invertebrates. Am. Malacol. Bull. 27, 25–45. https://doi.org/10.4003/006.027.0203

Williams, S.T., 2017. Molluscan shell colour. Biol. Rev. 92, 1039–1058. https://doi.org/10.1111/brv.12268

Wilson, L.J., Fulton, C.J., Hogg, A.M., Joyce, K.E., Radford, B.T.M., Fraser, C.I., 2016. Climate-driven changes to ocean circulation and their inferred impacts on marine dispersal patterns. Glob. Ecol. Biogeogr. 25, 923–939. https://doi.org/10.1111/geb.12456

Wolf, J.B.W., Ellegren, H., 2017. Making sense of genomic islands of differentiation in light of speciation. Nat. Rev. Genet. 18, 87–100. https://doi.org/10.1038/nrg.2016.133

WoRMS Editorial Board, 2021. World Register of Marine Species [WWW Document]. https://doi.org/10.14284/170

Wormuth, J.H., 1981. Vertical distributions and diel migrations of Euthecosomata in the northwest Sargasso Sea. Deep Sea Res. Part A, Oceanogr. Res. Pap. 28, 1493–1515. https://doi.org/10.1016/0198-0149(81)90094-7

Wormuth, J.H., 1985. The role of cold-core Gulf Stream rings in the temporal and spatial patterns of euthecosomatous pteropods. Deep Sea Res. Part A, Oceanogr. Res. Pap. 32, 773–788. https://doi.org/10.1016/0198-0149(85)90114-1

Wu, T.D., Watanabe, C.K., 2005. GMAP: A genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21, 1859–1875. https://doi.org/10.1093/bioinformatics/bti310

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., Chen, J., Chen, S., 2012. FastUniq: A fast *de novo* duplicates removal tool for paired short reads. PLoS One 7, e52249. https://doi.org/10.1371/journal.pone.0052249

Xuereb, A., Benestan, L., Normandeau, É., Daigle, R.M., Curtis, J.M.R., Bernatchez, L., Fortin, M.J., 2018a. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). Mol. Ecol. 27, 2347–2364. https://doi.org/10.1111/mec.14589

Xuereb, A., Kimber, C.M., Curtis, J.M.R., Bernatchez, L., Fortin, M.J., 2018b. Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. Mol. Ecol. 27, 5035–5048. https://doi.org/10.1111/mec.14942

Yang, H., Lohmann, G., Krebs-Kanzow, U., Ionita, M., Shi, X., Sidorenko, D., Gong, X., Chen, X., Gowan, E.J., 2020. Poleward shift of the major ocean gyres detected in a warming climate. Geophys. Res. Lett. 47, e2019GL085868. https://doi.org/10.1029/2019GL085868

Yeaman, S., Whitlock, M.C., 2011. The genetic architecture of adaptation under migration-selection balance. Evolution 65, 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Xi Ping Cuo, Z., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., Zheng, Hancheng, Liu, T., He, W., Li, K., Luo, R., Nie, X., Wu, H., Zhao, M., Cao, H., Zou, J., Shan, Y., Li, Shuzheng, Yang, Q., Asan, Ni, P., Tian, G., Xu, J., Liu, X., Jiang, T., Wu, R., Zhou, G., Tang, M., Qin, J., Wang, T., Feng, S., Li, G., Huasang, Luosang, J., Wang, W., Chen, F., Wang, Y., Zheng, X., Li, Z., Bianba, Z., Yang, G., Wang, X., Tang, S., Gao, G., Chen, Y., Luo, Z., Gusang, L., Cao, Z., Zhang, Q., Ouyang, W., Ren, X., Liang, H., Zheng, Huisong, Huang, Y., Li, J., Bolund, L., Kristiansen, K., Li, Y., Zhang, Y., Zhang, X., Li, R., Li, Songgang, Yang, H., Nielsen, R., Wang, Jun, Wang, Jian, 2010. Sequencing of fifty human exomes reveals adaptations to high altitude. Science 329, 75–78. https://doi.org/10.1126/science.1190371

Young, R.G., Abbott, C.L., Therriault, T.W., Adamowicz, S.J., Hogg, I., 2017. Barcode-based species delimitation in the marine realm: A test using Hexanauplia (Multicrustacea: Thecostraca and Copepoda). Genome 60, 169–182. https://doi.org/10.1139/gen-2015-0209

Yuasa, H., Kajitani, R., Nakamura, Y., Takahashi, K., Okuno, M., Kobayashi, F., Shinoda, T., Toyoda, A., Suzuki, Y., Thongtham, N., Forsman, Z., Bronstein, O., Seveso, D., Montalbetti, E., Taquet, C., Eyal, G., Yasuda, N., Itoh, T., 2021. Elucidation of the speciation history of three sister species of crown-of-thorns starfish (*Acanthaster* spp.) based on genomic analysis. DNA Res. 28, dsab012. https://doi.org/10.1093/dnares/dsab012

Zapata, F., Wilson, N.G., Howison, M., Andrade, S.C.S., Jörger, K.M., Schrödl, M., Goetz, F.E., Giribet, G., Dunn, C.W., 2015. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. Proc. R. Soc. B Biol. Sci. 282, 20141739. https://doi.org/10.1098/rspb.2014.1739

Zeebe, R.E., Ridgwell, A., Zachos, J.C., 2016. Anthropogenic carbon release rate unprecedented during the past 66 million years. Nat. Geosci. 9, 325–329. https://doi.org/10.1038/ngeo2681

Zelditch, M.L., Swiderski, D.L., Sheets, H.D., Fink, W.L., 2004. Geometric morphometrics for biologists: A primer. Elsevier Academic Press, San Diego and London.

Zieritz, A., Hoffman, J.I., Amos, W., Aldridge, D.C., 2010. Phenotypic plasticity and genetic isolation-by-distance in the freshwater mussel *Unio pictorum* (Mollusca: Unionoida). Evol. Ecol. 24, 923–938. https://doi.org/10.1007/s10682-009-9350-0

Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L., Yorke, J.A., 2013. The MaSuRCA genome assembler. Bioinformatics 29, 2669–2677. https://doi.org/10.1093/bioinformatics/btt476

Zouros, E., Ball, A.O., Saavedra, C., Freeman, K.R., 1994. An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. Proc. Natl. Acad. Sci. U. S. A. 91, 7463–7467. https://doi.org/10.1073/pnas.91.16.7463

Summary

Rising atmospheric CO_2 concentrations cause ocean acidification, a decrease in pH that threatens the existence of many marine calcifying organisms. Shelled pteropods are marine planktonic snails that are regarded as bioindicators of ocean acidification because their thin aragonitic shells are susceptible to dissolution. Despite their small body size, shelled pteropods play an important role in the ocean's carbonate budget and in marine food webs worldwide. Experiments have been conducted on shelled pteropods to assess their short-term responses to ocean acidification, but little is known about their evolutionary potential to adapt to long-term environmental changes. It is not possible to directly observe the evolutionary process due to difficulties in maintaining pteropods in lab cultures and their relatively long generation times. However, we can gain insight into factors affecting their adaptive potential by analysing levels of standing genetic variation within populations, gene flow between populations, and demographic fluctuations during their evolutionary history from genomic data.

While pteropods live in an open ocean environment and are hypothesised to have high effective population sizes and dispersal potentials, it is unknown if pteropod species are genetically homogeneous across their broad spatial ranges, or composed of several distinct populations. Given their roles as bioindicators, it is necessary to accurately assess their species boundaries because different species have different evolutionary trajectories and may have different sensitivities. In this thesis, I aimed to assess the spatial distribution of genetic variation within the shelled pteropod genus *Limacina*, to gain insight into the drivers of population structure in the open ocean and to obtain a better understanding of their evolutionary history and adaptive potential. *Limacina bulimoides* was chosen as a focal species because of its broad subtropical distribution and high abundance across the globe. Hence, levels of genetic variability in this species could be assessed at a population level across various spatial scales.

In **Chapter 2**, we assessed the population structure of *L. bulimoides* across a latitudinal transect in the Atlantic Ocean using partial DNA sequences of two barcoding genes, namely the mitochondrial cytochrome oxidase I (COI) and nuclear ribosomal 28S genes. Genetic differentiation of *L. bulimoides* across the sampling sites was compared to their shell shape variation and placed within the context of their abundance along an equivalent transect sampled two years later. We uncovered two dispersal barriers, one across the equatorial upwelling region between 15°N and 4°S, supported only by differentiation at the nuclear 28S locus, and the other dispersal barrier in the southern subtropical gyre, at 15-18°S, which was supported by both barcoding genes and shell shape variation. The locations of these dispersal barriers were congruent with regions of low abundance, supporting the hypothesis that areas of suboptimal habitat may function as barriers to dispersal in holoplanktonic organisms.

In **Chapter 3**, we developed a target capture approach to investigate genome-wide variation in the pteropod *L. bulimoides,* which was also tested on related pteropod species *L. trochiformis*, *L. lesueurii*, *L. helicina* and *Heliconoides inflatus*. A 2.9 gigabase draft genome of *L. bulimoides* was generated, and used in conjunction with a draft transcriptome to develop a set of genome-wide target capture probes, comprising 2812 single copy nuclear genes, including conserved protein coding regions, the 28S rDNA sequence, ten mitochondrial genes, 35 candidate biomineralisation genes and 41 non-coding regions. These probes were successful in obtaining detailed genomic information from the target species *L. bulimoide*s with 97% of the targets being recovered.

In **Chapter 4**, we applied the target capture probes developed in Chapter 3 to analyse spatial patterns of divergence of *L. bulimoides* across the global ocean. Genomic variation was studied with 107,214 single nucleotide polymorphisms (SNPs) from across 161 individuals, while shell shape variation was analysed using geometric morphometric analyses of shell images. We identified three distinct lineages, which we called the Atlantic, Indo-Pacific and Pacific lineage, based on their geography. We found no evidence of recent gene flow between the three lineages, not even between the Indo-Pacific and Pacific lineages that occur sympatrically in the North Pacific. The timing of divergence between the lineages was estimated to be during the mid-Pleistocene transition around 1 million years ago, while the fluctuations in population size within lineages coincided with known glacial-interglacial transitions. Shell shape was subtly different but overlapping between the lineages, and could not be used to distinguish them. However, we identified tissue pigmentation within the North Pacific individuals of the Pacific lineage as a potential distinguishing trait from the sympatric Indo-Pacific lineage. Hence, the circumglobal *L. bulimoides* is actually composed of three reproductively isolated lineages with more restricted distribution patterns that partially overlap.

In **Chapter 5,** we focused on the genome-wide diversity of 142 *L. bulimoides* individuals of the Atlantic lineage, with 97,425 SNPs obtained using target capture probes. The three populations that were tentatively identified in Chapter 2 were confirmed by the genome-wide analysis, namely the North, Equatorial and South Atlantic populations, with no evidence of recent gene flow between them. The dispersal barriers between the three populations were narrowed down to 14-15°N and 15-18°S. The presence of narrow dispersal barriers and absence of genetic mixing suggests that (bio-)physical barriers, natural selection, or a combination of both could be keeping populations apart, although more analyses are required to identify the processes maintaining this population structure. The mitochondrial and nuclear signals were incongruent for some individuals, which suggest (ancient) mitochondrial introgression between populations.

In **Chapter 6**, we demonstrated a new method in which alcohol-based handgel, which was widely distributed during the COVID-19 crisis, was used to position shelled pteropods under the microscope for standardised photographs and mor-

phometric analysis. The new method was more efficient than previous positioning methods used. There is potential for broader application of this method for the taxonomic identification, and the morphological and ontogenetic study of other small molluscs and planktonic organisms.

In **Chapter 7**, I summed up the findings from Chapters 2-6 and point to future research directions. My thesis has shown that genome-wide markers provide additional insights into the population structure and evolutionary history of *L. bulimoides* compared to barcoding genes. I found that *L. bulimoides* is not genetically homogeneous across its range, but is composed of at least three reproductively isolated lineages across the Atlantic, Indian and Pacific Oceans. Detailed analysis of the Atlantic lineage revealed further population structure, with three distinct populations separated by narrow dispersal barriers. Looking forward, the methods used to access genomic information in *L. bulimoides* can be applied to other shelled pteropods, including species in (sub)polar regions, which are already experiencing the effects of a rapidly acidifying ocean.

Samenvatting

De toenemende hoeveelheid CO_2 die door menselijke activiteiten in de atmosfeer wordt gebracht leidt tot verzuring van de oceaan. Dit is een proces waarbij het zeewater geleidelijk zuurder wordt, waardoor het voor kalkvormende organismen steeds moeilijker wordt om hun schaaltjes of schelpen te bouwen. Pteropoden, ook wel zeevlinders genoemd, zijn een groep planktonslakken die extreem gevoelig lijken te zijn voor oceaanverzuring vanwege hun dunne huisjes gemaakt van aragoniet (een zeer oplosbare vorm van calciumcarbonaat). Ook al zijn zeevlinders slechts enkele millimeters tot een centimeter groot, ze spelen een belangrijke rol in de mariene voedselketen en het carbonaat budget van de oceaan. Korte termijn experimenten hebben aangetoond dat zeevlinders te lijden hebben onder oceaanverzuring. Echter, we weten niet of zeevlinders zich kunnen aanpassen op de lange termijn. Het is niet mogelijk om het evolutionaire proces direct te observeren want het is zeer moeilijk om zeevlinders in het lab te kweken en ze hebben een relatief lange generatietijd (ongeveer een jaar). Maar we kunnen wel inzicht krijgen in factoren die hun aanpassing op de lange termijn bepalen, zoals de hoeveelheid genetische variatie binnen en tussen populaties, de mate van uitwisseling tussen populaties, en fluctuaties in populatiegrootte gedurende hun evolutionaire geschiedenis. Dit kan worden afgeleid uit het DNA.

Pteropoden leven in de open oceaan en algemeen wordt aangenomen dat ze grote effectieve populaties hebben en zich zeer goed kunnen verspreiden. Echter, we weten niet of soorten genetisch homogeen zijn over het hele gebied waarin ze voorkomen, of dat ze uit verschillende populaties of (onder)soorten bestaan. Omdat zeevlinders worden gezien als bio-indicatoren of graadmeters van oceaanverzuring is het belangrijk om de diversiteit aan soorten goed te kennen. Verschillende soorten hebben namelijk een verschillende evolutionaire geschiedenis en kunnen daarom verschillend reageren op veranderingen in hun leefomgeving. In dit proefschrift heb ik gekeken naar de verspreiding van genetische variatie in het genus *Limacina* om meer inzicht te krijgen in de processen die een rol spelen in het ontstaan van verschillende populaties en soorten in de open oceaan, alsmede hun potentieel om zich te kunnen aanpassen aan een toekomstige oceaan. Mijn onderzoek heeft zich met name gericht op *Limacina bulimoides* omdat deze soort in grote aantallen voorkomt en een wereldwijde verspreiding heeft. Hierdoor kon ik de genetische variatie goed bestuderen aan de hand van voldoende individuen uit verschillende oceaanbekkens.

In **Hoofstuk 2** hebben we de populatiestructuur van *L. bulimoides* bestudeerd langs een noord-zuid transect in de Atlantische Oceaan aan de hand van twee barcoding genen, namelijk het mitochondriale cytochroom oxidase I (COI) gen en het nucleaire ribosomale 28S gen. We hebben genetische differentiatie tussen populaties vergeleken met morfologische variatie in schelpvorm, en ook gekeken naar de abundantie langs een vergelijkbaar transect dat twee jaar later is bemonsterd.

We ontdekten twee dispersie barrières: één barrière was gelegen in het gebied van de equatoriale oceaanstromen tussen 15°N en 4°S en werd alleen ondersteund door het nucleaire 28S gen. De tweede barrière was in de Zuid-Atlantische gyre tussen 15°S en 18°S, en werd ondersteund door beide genen en door verschillen in schelpvorm. De locaties van deze barrières overlapten met een afname in populatie dichtheden. Dit ondersteunt de hypothese dat suboptimale gebieden mogelijk leiden tot dispersie barrières in het plankton van de open oceaan.

In **Hoofdstuk 3** hebben we een 'target capture' methode ontwikkeld om variatie verspreid over het hele genoom van de pteropode *L. bulimoides* te onderzoeken. Daarnaast hebben we deze methode getest op de verwante soorten *L. trochiformis*, *L. lesueurii*, *L. helicina* en *Heliconoides inflatus*. Op basis van een eerste gefragmenteerd genoom van 2.9 gigabase en een transcriptoom van *L. bulimoides* zijn de capture probes ontwikkeld. Deze probes omvatten 2.812 'single copy' nucleaire genen, het 28S rDNA gen, tien mitochondriale genen, 35 kandidaat biomineralisatie genen en 41 niet-coderende stukken. De probes waren succesvol om variatie over het hele genoom van *L. bulimoides* te onderzoeken (97% van de beoogde DNA sequenties zijn teruggevonden).

In **Hoofdstuk 4** zijn de 'target capture probes', die we hebben ontwikkeld in Hoofdstuk 3, gebruikt om de ruimtelijke verspreiding van genoom variatie in *L. bulimoides* over de wereldwijde oceaan te onderzoeken. Genoom variatie is bestudeerd aan de hand van 107.214 'single nucleotide polymorphisms' (SNPs) van 161 individuen, terwijl hun schelpvorm werd onderzocht met een geometrische morfometrische analyse op basis van gestandardiseerde foto's. We ontdekten drie evolutionaire lijnen van verwantschap, die we de Atlantische, Indo-Pacifische en Pacifische lijn hebben genoemd op basis van hun verspreidingen. We vonden geen enkel bewijs van genetische uitwisseling tussen deze drie lijnen, zelfs niet tussen de Indo-Pacifische en Pacifische takken welke samen voorkomen in de Noord Pacifische Oceaan. Op basis van genetische differentiatie schatten we dat de lijnen zijn gesplitst tijdens het midden-Pleistoceen, ongeveer 1 miljoen jaar geleden. Fluctuaties in populatiegrootte van de lijnen lijken samen te vallen met transities tussen ijstijden en interglacialen. De schelpvorm verschilde subtiel tussen de verschillende lijnen maar vertoonde ook overlap, en kan daarom niet gebruikt worden als een morfologisch kenmerk. Echter, we hebben pigmentvlekken op de 'vleugels' van pteropoden van de Pacifische lijn ontdekt, die een mogelijk kenmerk zijn om de Pacifische lijn van de sympatrische Indo-Pacifische lijn te kunnen onderscheiden. Dus hoewel *L. bulimoides* een wereldwijde verspreiding heeft, bestaat deze 'soort' in feite uit drie reproductief geïsoleerde lijnen (wellicht zelfs drie afzonderlijke soorten) met meer beperkte en deels overlappende verspreidingspatronen.

In **Hoofdstuk 5** hebben we de genoom-wijde variatie van 142 *L. bulimoides* individuen uit de Atlantische Oceaan onderzocht, op basis van 97.425 SNPs verkregen middels de 'target capture probes' uit Hoofdstuk 3. De resultaten van deze genoomwijde studie bevestigen dat de Atlantische lijn bestaat uit drie verschillende popu-

laties, zoals eerder voorgesteld in HOOFDSTUK 2, namelijk een Noord-Atlantische, Equatoriale, en Zuid-Atlantische populatie, zonder enige indicatie van recente genetische uitwisseling tussen deze populaties. De twee dispersie barrières zijn nader gespecificeerd en gelokaliseerd op 14-15°N en 15-18°S. Het bestaan van dergelijke nauwe dispersie barrières in de open oceaan en het ontbreken van genetische uitwisseling, suggereert dat fysische barrières zoals oceaanstromingen, natuurlijke selectie, of een combinatie hiervan heeft geleid tot het uit elkaar houden van de populaties. Echter, verdere analyses zijn nodig om meer inzicht te krijgen in de processes die verantwoordelijk zijn voor deze populatie structuur. De mitochondriale en nucleaire signalen waren soms tegenstrijdig, wat betekent dat er waarschijnlijk (historische) introgressie tussen de populaties heeft plaatsgevonden.

In **HOOFDSTUK 6** beschrijven we een nieuwe methode, gebruikmakend van de alcohol-handgel die ruim voor handen was tijdens de COVID-19 crisis, om zeevlinders te positioneren onder de microscoop en gestandardiseerde foto's te maken voor morfometrische analyses. De nieuwe methode bleek veel efficiënter dan de eerder gebruikte methodes. Daarnaast zou deze methode kunnen worden toegepast voor morfologische identificatie en ontogenetische studies van andere kleine mollusken en plankton.

In **HOOFDSTUK 7** heb ik alle bevindingen van HOOFDSTUKKEN 2-6 samengevat en richtingen voor toekomstig onderzoek aangegeven. Mijn proefschrift heeft laten zien dat genoom-wijde merkers meer inzicht geven in de populatiestructuur en evolutionaire geschiedenis van *L. bulimoides* dan de barcoding genen die gewoonlijk gebruikt worden. Ik heb gevonden dat *L. bulimoides* niet genetisch homogeen is over het hele verspreidingsgebied, maar bestaat uit tenminste drie genetisch geisoleerde evolutionaire lijnen in de Atlantische, Indische en Pacifische Oceaan. Binnen de Atlantische lijn vonden we nog verdere ruimtelijke structuur met drie populaties gescheiden middels twee nauwe dispersie barrières. Kijkend naar de toekomst, zie ik dat de nieuwe methoden die ik heb gebruikt om te kijken naar variatie binnen het genoom van *L. bulimoides* ook kunnen worden toepast op andere zeevlinders, wat met name interessant zal kunnen zijn voor de (sub)polaire pteropode soorten die op dit moment al te lijden hebben onder de steeds zuurder wordende oceaan.

Author contributions

Chapter 2

LQC, TMPB, EG and KTCAP contributed to the study design. EG and KTCAP collected samples for the study. LQC and TMPB collected molecular and morphometric data. All authors analysed the data. LQC wrote the manuscript, with input from all authors. All authors approved of the final manuscript.

Chapter 3

LQC, TMPB, MC, MK, IS, GH and KTCAP contributed to the study design. KTCAP provided samples used in this study. FM and PRS analysed sequence data and contributed to the capture design. LQC, TMPB, MK and IS contributed to the molecular work. LQC, TMPB, MC, GH and KTCAP contributed to bioinformatic analyses and manuscript writing. All authors provided feedback and approved of the final manuscript.

Chapter 4

LQC, GH and KTCAP designed the study. EG and KTCAP contributed samples used in this study. LQC collected the molecular data while GS collected the morphometric data. LQC, MC, GH and KTCAP contributed to the bioinformatic analyses while LQC and GS analysed the morphometric data. LQC, MC, GH, EG and KTCAP contributed to the manuscript writing. All authors provided feedback and approved of the manuscript.

Chapter 5

LQC, GH and KTCAP designed the study. LQC, EG and KTCAP collected samples for the study. LQC collected the molecular data, while GS collected the morphometric data. LQC, GS, GH, and KTCAP contributed to data analyses and LQC, GH, JH and KTCAP contributed to manuscript writing. All authors provided feedback and approved of the manuscript.

Chapter 6

LQC and KTCAP designed the study. GS processed the specimens for photography, micro-CT scanning and DNA barcoding. LQC wrote the manuscript and all authors provided feedback and approved of the final manuscript.

About the author

Le Qin was born on 2nd September 1993 in Singapore, where she spent most weekends of her childhood exploring the beaches and various nature reserves. This is where her love for nature and science all began. She got her first taste of research in junior college, when she participated in the Science Research Programme, investigating the distribution and diets of intertidal limpets at the Tropical Marine Science Institute (TMSI) at the National University of Singapore (NUS). After graduating from junior college, Le Qin continued at TMSI, this time to study the formation of hairs on some marine mussel species, together with her twin sister. She was awarded a Loke Cheng-Kim scholarship to pursue a Bachelor's degree in Zoology at Fitzwilliam College, University of Cambridge, England from 2012 to 2015. There were many opportunities for fieldwork and experiments, including studying the secretions of dock beetle larvae and habitat choice in estuarine shrimp, and from these Le Qin decided that working with live animals was probably not so suitable for her. A volunteering stint at the Mollusc collections of the Natural History Museum in London further solidified her view that dead animals were much more cooperative. Seeking a return to her marine roots, she wrote her final year thesis on the relative importance of various habitats to commercially important demersal fish, and the case for including these habitats in marine protected areas.

Le Qin then continued with a Master's degree in Biosystematics based at the Natural History Museum and Imperial College London from 2015 to 2016, where she completed three research projects in a diverse range of fields, including fly metagenomics, tissue visualisation techniques for tapeworms, and a project on the biogeography of riverine snails, supervised by Ellinor Michel and Jon Todd. Following the completion of her Master's degree, Le Qin returned to Singapore and kept busy with sorting out deep sea benthic organisms at the Keppel Corporate Lab (NUS), while looking out for exciting PhD opportunities. As fate would have it, Ellinor and Jon forwarded her an email about the opening for this current PhD position at Naturalis Biodiversity Center, and she seized her chance. In October 2017, she started her PhD project in Leiden on the population structure of shelled pteropods, under the supervision of Katja Peijnenburg and Galice Hoarau. There were many adventures along the way, such as living in the northernmost country she'd ever been to (which was not as cold or terrifying as she imagined) and embarking on a 25-day research cruise across the Atlantic Ocean to collect plankton for her genomic analyses. With her work, she gained greater insight into the distribution of genetic variation in pelagic snails, and facilitated the application of genome-wide techniques in non-model planktonic organisms. Le Qin will join the lab of Roger Butlin in April this year to continue working on snails, this time on the role of chromosomal inversions in the adaptation and speciation of intertidal periwinkles.

Peer reviewed publications

Choo, L.Q., Bal, T.M.P., Goetze, E., Peijnenburg, K.T.C.A., 2021. Oceanic dispersal barriers in a holoplanktonic gastropod. J. Evol. Biol. 34, 224–240.

Choo, L.Q., Bal, T.M.P., Choquet, M., Smolina, I., Ramos-Silva, P., Marlétaz, F., Kopp, M., Hoarau, G., Peijnenburg, K.T.C.A., 2020. Novel genomic resources for shelled pteropods: A draft genome and target capture probes for *Limacina bulimoides*, tested for cross-species relevance. BMC Genomics 21, 11.

Wall-Palmer, D., Janssen, A., Goetze, E., Choo, L.Q., Mekkes, L., Peijnenburg, K.T.C.A., 2020. Fossil-calibrated molecular phylogeny of atlantid heteropods (Gastropoda, Pterotracheoidea). BMC Evol. Biol. 20, 124.

Macher, J.N., Speksnijder, A., Choo, L.Q., van der Hoorn, B., Renema, W., 2019. Uncovering bacterial and functional diversity in macroinvertebrate mitochondrial-metagenomic datasets by differential centrifugation. Sci. Rep. 9, 10257.

Choo, L.Q., Crampton-Platt, A., Vogler, A.P., 2017. Shotgun mitogenomics across body size classes in a local assemblage of tropical Diptera: Phylogeny, species diversity and mitochondrial abundance spectrum. Mol. Ecol. 26, 5086–5098.

Choo, L.M., Choo, L.Q., Tan, K.S., 2014. The origin and formation of hair on external valve surfaces of the tropical marine mussel *Modiolus traillii* (Reeve, 1857). J. Molluscan Stud. 80, 111–116.

Acknowledgements

This thesis would not have been possible without the help of many people.

Firstly, I'd like to express my deepest appreciation to my supervising committee, **Katja**, **Galice** and **Jef**.

I am extremely grateful to **Katja** for giving me the opportunity to work with pteropods. I had never seen such delicate and exquisite molluscs before starting my PhD, and it was amazing to see them alive in the field! I've learnt a lot from you about being a scientist, juggling workloads with grace and inspiring others through your teaching and outreach. Thank you for the many opportunities to attend international conferences and courses, and expand my scientific network and skills in population genomics.

Galice, thank you for your unwavering support and guidance and for helping to keep me on track with deadlines. I will fondly remember the precious time I spent in Bodø, surrounded by a welcoming and supportive academic community, and the amazing Norwegian nature and horses!

Thank you, **Jef**, for your invaluable experience and advice with navigating the entire PhD process. I greatly appreciated your guidance in putting the thesis together into its final format, as well as making the final steps as smooth as possible!

I would also like to extend my gratitude to my exam committee: **Prof Kerstin Johannesson**, **Prof Menno Schilthuizen**, **Dr Lisa Becking**, **Prof Willem Renema**, **Prof Astrid Groot**, **Dr Susanne Wilken** and **Prof Corina Brussaard** for taking the time out of their busy schedules to read my thesis and participate in my defence.

I am also greatly indebted to the plankton team.

Thank you **Debbie** and **Paula**, for your pteropod expertise and advice about pursuing a research career, as well as numerous cat sitting opportunities that I thoroughly enjoyed. I count myself as so lucky to be able to learn from you. **Daniela**, **Catharina**, **Giada and Pane**, thank you for the wonderful opportunity to supervise your work, and for your trust in me. A special mention to **Giada**, whose work appears in several chapters within this thesis, your enthusiasm and thirst for knowledge has kept me motivated even in the darkest of times.

Lisette, my dearest slakzus, thank you for being my PhD buddy and emotional support. I don't think I could have completed this journey without you walking by my side. I've enjoyed our time together, including the authentic Dutch Easter and Christmas with your family, our lovely plankton-filled time on Leg 8 abroad the Pelagia and getting lost in the crowd in a random pub in Ireland.

Thijs, thank you for being the best lab partner! You welcomed me into your home when I first arrived in the Netherlands, then we set off to Bodø together, where we spent countless hours in the lab extracting DNA and preparing libraries. Along the way, I picked up Dutch vocabulary and grammar, while you got yourself

a new PhD project there (among other stuff;)). It was a pleasure to work with you, and continue from where your Masters project left off.

Bo, Boris, Gloria, Iliana, Jaap, Luis, Mari-Lee, and other plankton team members, it was great sharing such pterrific pteropody times with you all.

And to the late **Arie Janssen**, thank you for sharing with us your valuable pteropod insights and knowledge. It was a real privilege to have known and worked with you.

I'd also like to extend my sincere thanks to my dear colleagues in Bodø.

Marvin, thank you for your expertise with the target capture technique, pipeline and analyses. I managed to have a running start with my project because of this, and I am very grateful for your dedication to your work, and the fun and spontaneous times we shared outside of the lab!

This project also would not have been possible if not for **Martina**. Thank you for teaching me about all the various lab techniques, and for keeping the labs so well-organised. It is always a successful lab day when working with you!

Apollo, thank you for being my lab buddy to liven up the long hours spent there. I've enjoyed our time together, collaborating on analyses and having adventures in Svalbard and awesome gatherings with great food.

Alexander and **Lars Martin**, thank you for managing the server resources and ensuring that they ran smoothly, these computers were certainly indispensable for all my genomic analyses!

I'd also like to thank my other colleagues who provided their expertise and insight: **Irina**, **Aurélien**, **Isabel**, **Mads** and **Leslie**, as well as the other scientific and administration staff at Nord University who've helped me along the way and looked out for me during my time in Bodø.

To **Erica Goetze**, thank you for the opportunity to work together, I have learnt a lot from your precise insight in the field of biological oceanography. It was great to meet in person at OSM2020, and I'm looking forward to our next (virtual) meeting at OSM2022!

And to **Ferdi Marlétaz**, even though we may not have met in person (yet), I've greatly enjoyed our collaborations on the pteropod target capture, and thank you for your advice and helpful contributions in our ongoing bioinformatics efforts!

I also gratefully acknowledge the numerous plankton collections that my work depended on, and thank **Alice Burridge**, **Michelle Jungbluth** and **Atsushi Tsuda** for collecting these precious samples and kindly making them available.

Wayne, my fellow bbt and chicken nuggets fiend, and token physicist friend, thank you for your help in proofreading of my thesis sections. Your advice and encour-

agement from the perspective of someone who recently completed their PhD was irreplaceable. I look forward to more deep conversations about science and navigating the academic world, as well as on other less professional topics.

I am also grateful towards the members of the Marine biodiversity group: **Willem**, **Nicole**, **Frank**, **José**, **Aleks**, **Eduard**, **Esther**, **Jan**, **Martina,** S**abrina**, **Sander**, **Thomas**, and other recent members, for the interesting and fruitful scientific discussions, cake and lovely lab retreats. I had so much fun on our trip to the Biesbosch, and our ice-skating trip in Kinderdijk! I'd like to especially thank **Willem** for making the group a wonderful place for research, and your helpful insight about fossils and geological timescales. And thank you, **Jan**, for the opportunity to collaborate on a project together!

Thank you to the laboratory staff at Naturalis: **Arjen**, **Elza**, **Bertie-Joan**, **Dirk**, **Frank**, **Kees**, **Marcel**, **Marina, Rob** and **Roland** for their technical support and for facilitating my experiments, as well as the front-desk, **Angela** and **Anneke**, for helping me with the most random queries and being a warm familiar face. I also appreciate the help and guidance from the Naturalis RCO, including **Dominique**, **Erik**, and the other members of the **PhD council** for making the Naturalis a wonderful and supportive place for us.

Geert-Jan Brummer, thank you for the opportunity to rummage through your plankton samples at the VU, and for your advice and encouragement! It has meant a lot to me.

To the lovely people on board the Pelagia for NICO expedition leg 8, thank you for making my first scientific cruise experience so great! Thank you to all the scientific staff and crew for patiently showing me the ropes as I grew my sea legs, including our cruise leader **Corina** for ensuring everything was well-organised, from sampling locations to Pelagia's-got-Talent and **Susanne** for being a supportive and encouraging presence for me when I was adjusting to life on board. A special mention to **Martimartin fan club** for the fun and laughter we shared.

Milan Malinsky, **Emiliano Trucchi** and **Michael Matschiner**, thank you for your technical expertise and advice during and after the population genomics course in Cesky Krumlov. I've learnt a lot from all of you and your inspiring work.

To the rest of my scientific network whom I've met along the way: **Paula Pappalardo**, **Leocadio Blanco-Bercial** and **Amy Maas**, **Katie Peichel** and **Sean Stankowski**, thank you for the encouragement and helpful advice for my scientific career.

I'd also like to recognise the assistance I received from UvA and IBED, from the staff who took an interest in my work to the opportunities to be part of a wider academic community, and the UvA Care project, from which I benefited greatly.

Special thanks to **Jan Bruin** for the layout and formatting of the thesis and **Amalia Aikaterini Mailli** and **George Komiotis** for bringing my cover design idea to reality.

The spicy girls: **Ajaree**, **Dewi**, **Deyi**, **Eka**, **Richa**, **Nida**, thank you for being my go-to bunch of friends when things get tough! I am so grateful for all the wonderful moments we shared (usually involving delicious food and typical Dutch tourist attractions). To the Speedo body club: **Andrés**, **Esther**, **Héctor**, **Lisette**, **Lizzie**, and other guest stars, thank you for keeping me sane and healthy during the pandemic. Many thanks also to the members of the several in-person and online game groups, including **Andrés**, **Charlotte**, **Eka**, **Esther**, **Héctor**, **Kevin**, **James**, **Lizzie**, **Manon**, **Merijn**, **Panos, Roderick**, and **Werner**, we now know who is great at drawing, who is a terrible liar, and who can (probably) get away with murder. To those people above, **Bob**, **Giada**, **Gloria**, **Izai**, **Kasper, Marcel**, **Pane**, **Zac**, and the other masters and bachelor students in the PhD room, thank you for the wonderful company, camaraderie and cake that we shared.

To my Singaporean community in the Netherlands: **Agnes**, **Angie**, **Evelyn**, **Doreen** and **Maarten**, **Mickey**, **Grace**, **Sheena**, and **Wayne**, thank you for all the wonderful gatherings and great food, it definitely alleviated my bouts of homesickness!

To the PhD and international student community in Nord University, thank you for welcoming me into your midst multiple times, we've had such awesome adventures and the most amazing parties, both under the midnight sun and the northern lights. To my friends: **Adriána**, **Amalia**, **Ana**, **Alex**, **Artem**, **Apollo**, **Arun**, **Aurélien**, **Betty**, **Bingqian**, **Chen**, **Chris**, **Deepti**, **Dhurba**, **Èric**, **Fernando**, **Ingrid, Isabel**, **Isabelle**, **Joost**, **Joshua**, **Julia**, **Leona**, **Liu Cui**, **Marvin**, **Park**, **Paula**, **Prabhu**, **Rabbani**, **Sanne**, **Shuhua**, **Spoorti**, **Thijs**, **Tiago**, **Ying**, **Yousri**, **Zhiyuan**, **Ørjan**, and any others I may have inadvertently missed, I've thoroughly enjoyed the times I spent hiking/fishing/cooking/eating/skiing/ice-skating/horse-riding with you all, and there will always be a special place in my heart for you. **Aurélia** and **Hin**, thank you for the lovely horsey times we shared in Valnesfjord. And it is so special to find a fellow Singaporean in a small city in Norway, thank you **Grace** and the **Coldevin family** for the numerous dinners and teas, weekend hikes, and for being a warm welcoming home in the midst of Bodø's dark and rainy days.

Many other people have inspired me and nurtured my interest in science, by showing me that it was fun, before anything else. Special mention to **Dr Tan Koh Siang**, who mentored me when I was 16 and excited about plucking limpets off rocks to dissect them. I still remember the hawker food analogies you used to explain diet choice to me, and I was really fascinated when I realised that I could, in fact, make studying snails my career choice. I am also very grateful towards **Ellinor** and **Jon**, with whom I had the wonderful opportunity to study river snails. Thank you for your guidance and confidence in me, and for sending information about this PhD opportunity my way! I'd also like to thank **David Reid**, **Andreia Salvador**, **Jon Ablett**, **John Taylor** and others from the NHM Mollusc department for indulging in

my obsession about snaily things, and giving me an opportunity to volunteer in the collections when I was just an undergrad. I also wish to extend my gratitude to the **Loke Cheng Kim Foundation**, which generously sponsored my undergraduate studies at Fitzwilliam College, University of Cambridge. I have been very privileged to explore so many opportunities and build my scientific career because of this, and I hope that I will be able to pass it forward and give back to the community in the future.

Finally, I wish to express my deepest gratitude to my dear **parents** and **sister** for their unconditional support for my career choice, and for letting me know that I am always loved despite the large physical distances between us. Thank you for your hard work and sacrifice through the years, and I love you all too.