Genomic variability and population structure in shelled pteropods

Choo, L.Q.

Publication date
2022

Citation for published version (APA):
1

Introduction
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-
Introduction

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 reproductively 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 adaptation 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 clymena) is a hybrid species of two putative parental species, Stenella coeruleoalba and Stenella longirostris, all three of which occur in sympathy. 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).
Chapter 1

Ecological 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; Biene 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 cryptis, 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 CO₂ 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
Introduction

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; García 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).
Chapter 1

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
Introduction

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 the cosomes 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
Introduction

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

**Limacina** species are characterised by their thin, aragonitic and sinistral 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...
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-
Introduction

*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 m³. A similar transect from three years later (in 2017) resulted in a maximum abundance of ~350 individuals per 1000 m³ 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](image-url)
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 m³. *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 analysis. 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, 2021).

**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 m³ 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 m³ with a maximum of
Introduction

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.
Chapter 1

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 Breilsford, 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. bullimoides*.

**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 (Fiser 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
Introduction

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. bulimoides*, 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 1

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.