Phylogeography of amphi-boreal marine fauna

HANNA LAAKKONEN

Finnish Museum of Natural History
and
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki
Finland

Academic Dissertation
To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in the Auditorium 116, Yrjö-Koskisen katu 1, Helsinki, on November 27th 2015 at 12 noon.

HELSINKI 2015
CONTENTS

Contents ........................................................................................................................................... 3
List of Publications .......................................................................................................................... 5
Abstract ........................................................................................................................................... 6
SUMMARY ........................................................................................................................................... 7

1. Introduction .................................................................................................................................. 7

1.1. Biogeographical framework ................................................................................................. 7

1.1.1. Phylogeography ................................................................................................................. 7

1.1.2. Amphi-boreal marine fauna ............................................................................................. 7

1.1.3. Paleogeography and climatic history ................................................................................. 8

1.2. Phylogeographical and systematic background ................................................................... 9

1.2.1. Phylogeographical patterns in amphi-boreal taxa ......................................................... 9

1.2.2. Genetic introgression ...................................................................................................... 10

1.2.3. Cryptic systematic diversity in marine fauna ................................................................. 11

1.3. Study systems ....................................................................................................................... 11

1.3.1. Pacific and Atlantic herrings – an example of amphi-boreal species pair ....................... 11

1.3.2. The near-cosmopolitan bivalve Hiatella arctica ............................................................. 12

2. Objectives .................................................................................................................................... 13

3. Material and Methods .............................................................................................................. 15

3.1. Material .................................................................................................................................. 15

3.2. Molecular markers ............................................................................................................... 15

3.3. Data analyses ....................................................................................................................... 17

3.3.1. Molecular diversity and phylogenetic analyses ............................................................... 17

3.3.2. Intraspecific patterns and recent trans-Arctic connections ............................................. 17

3.3.3. Substitution rates and divergence time .......................................................................... 18

3.3.4. Taxon delimitation ......................................................................................................... 19

3.3.5. Coincidence of phylogeographical split times .............................................................. 19

3.3.6. Analyses of introgression ............................................................................................... 19

4. Results and Discussion ............................................................................................................. 20

4.1. The diversity of inter-oceanic genetic distances ................................................................. 20

4.2. Phylogeographical patterns in amphi-boreal taxa .............................................................. 22

4.2.1. Vicariant histories without further dispersal ................................................................. 22

4.2.2. Several trans-Arctic divergence events ....................................................................... 23

4.3. Cryptic amphi-boreal taxa ................................................................................................. 24

4.4. Post-glacial dispersal ............................................................................................................ 24
Box 1. Cryptic diversity in the *Hiatella* bivalve complex (III) ..................25

Box 2. Post-glacial invasion and secondary contacts in herrings
and in the *Hiatella* bivalve complex (I) .............................................26

4.5. Secondary contacts of sister taxa following re-invasion from the Pacific .27

Box 3. Introgression between the Atlantic and Pacific herrings ..........28

4.6. Introgression following secondary contacts ........................................29

4.7. Factors facilitating trans-Arctic invasion and colonization of new areas .29

5. Conclusions and Future Prospects ....................................................................32

Acknowledgements ..........................................................................................34

References ........................................................................................................36
LIST OF PUBLICATIONS

This thesis is based on the following original papers, which are referred to by their Roman numerals:


<table>
<thead>
<tr>
<th>Contributions</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original idea</td>
<td>RV</td>
<td>RV, HL, PS</td>
<td>RV</td>
<td>RV</td>
</tr>
<tr>
<td>Material (sampling)</td>
<td>DL, PS, HL, RV</td>
<td>HL, PS, DL, RV</td>
<td>HL, PS, RV</td>
<td>HL, MH, PS, RV</td>
</tr>
<tr>
<td>Data production</td>
<td>HL</td>
<td>HL, PS</td>
<td>HL</td>
<td>HL, MH</td>
</tr>
<tr>
<td>Data analyses</td>
<td>HL</td>
<td>HL, PS, RV</td>
<td>HL</td>
<td>HL, PS</td>
</tr>
<tr>
<td>Manuscript preparation</td>
<td>HL, RV</td>
<td>HL, RV, PS</td>
<td>HL, RV</td>
<td>HL, RV</td>
</tr>
</tbody>
</table>

HL = Hanna Laakkonen  
RV = Risto Väinölä  
PS = Petr Strelkov  
MH = Michael Hardman  
DL = Dmitry Lajus
ABSTRACT

The Northern Atlantic and Pacific Oceans share a faunal element consisting of pairs of closely related vicariant taxa, known as the amphi-boreal fauna. The inter-oceanic systematic affinities reflect a history of shared ancestry since dispersal through the Bering Strait and across the Arctic basin, which was initiated after the opening of the Bering Strait at the end of the Miocene. This thesis examines the biogeographical history of the amphi-boreal faunal element using information from molecular marker characters. The aim is to document the dynamics and consequences of the faunal interchange between the two oceans, spanning from the end of Miocene to the present. This is done by comparing differences in the mitochondrial gene sequence variation in different taxonomic groups and across the circum-boreal geographical scale, encompassing 70 taxa. The consequences of dispersal to genetic diversity was examined more closely and with additional markers in two exemplary genera, herrings and bivalve molluscs of the genus *Hiatella*.

The phylogeographical histories of the Pacific–Atlantic amphi-boreal taxa were found to be remarkably variable. A simple vicariant history with a single early dispersal event in Pliocene or Early Pleistocene was inferred only in about half of the examined taxa, whereas signs of more than one trans-Arctic dispersals were found in a third of the cases. A close inter-oceanic relationship that reflects recent trans-Arctic dispersal or ongoing gene flow was found to be common for the amphi-boreal genera (40% of examined taxa). Overall the inter-oceanic mitochondrial divergence estimates within each of the broader taxonomic groups (fishes, molluscs, crustaceans, echinoderms, polychaetes) varied greatly, up to 10–20 fold, and suggest that trans-Arctic faunal dispersal has been a repeated process through the considered time frame. The divergence estimates of invertebrate taxa detected several instances of putative new allopatric species, not previously recognised. In the nearly cosmopolitan bivalve mollusc *Hiatella*, altogether 11 cryptic taxa were discovered in the Northern Hemisphere and a hypothesis of their repeated inter-oceanic movements was presented.

The consequences of secondary dispersal from the Pacific to Atlantic were studied more closely in herrings. The Pacific herring *C. pallasii* invaded the NE European seas post-glacially and differentiated into separate regional populations. It came into secondary contact with the native sister species Atlantic herring *C. harengus* and despite ecological and distributional differences of these species, hybridization and introgression from the Atlantic to Pacific herring has occurred, to different amounts in different local populations. Particularly the Balsfjord herring stock was judged to represent a hybrid swarm with an equilibrium mixture of genetic components of the two species in 20:80 proportions. This kind of generation of new synthetic diversity through the contact of old temporarily isolated inter-oceanic lineages has parallels in analogous histories inferred for the bivalves *Macoma* and *Mytilus*. The introgression between old lineages, but also the effects of shorter isolation-dispersal cycles, may be important but so far overlooked phenomena in shaping the genomic composition and adaptive properties of taxa that inhabit the post-glacially colonized marginal habitats of the northern seas. The current global warming and changing navigation practices are expected to enhance the trans-Arctic inter-oceanic faunal exchange in near future and also expose the amphi-boreal taxa to further genetic influence from their sister lineages.
SUMMARY

1. INTRODUCTION

1.1. Biogeographical framework

1.1.1. Phylogeography
Phylogeography is the study of spatial and temporal distributions of genetic lineages, especially within and among closely related species (Avise 2000). It was initially introduced as a new branch of biogeography that could connect gene genealogies to geography and at the same time serve as a bridge between population genetics and systematics (Avise et al. 1987). Comparative phylogeographic assessments can reveal general patterns in the regional biogeographical history of co-distributed species assemblages (e.g. Bermingham and Avise 1986). The congruent or incongruent patterns detected among taxa are used to interpret spatial and temporal biotic events, as related to climatic history on one hand and to species-specific biological properties (such as dispersal abilities) on the other. In the context of boreal marine fauna, the climatic history of the Pliocene and Pleistocene glacial cycles has been an essential element in shaping the species phylogeographical patterns (e.g. Briggs 2003, Maggs et al. 2008).

1.1.2. Amphi-boreal marine fauna
The boreal faunas of the North Atlantic and North Pacific Oceans include numerous pairs of vicarious taxa or populations which are closely related despite the current geographical isolation. This biogeographical element, the amphi-boreal fauna, reflects a history of shared ancestry attributable to past dispersal through the Bering Strait and across the Arctic basin (Durham and Mac-Neil 1967). Molecular studies on particular amphi-boreal species or groups often concur with this supposed vicariance, but sometimes they have revealed unexpectedly complex genetic subdivisions and biogeographical relationships (e.g. McDonald et al. 1991, Ortí 1994, Coulson et al. 2006, Nikula et al. 2007, Ursvik et al. 2007, Dodson et al. 2008, Durbin et al. 2008, Carr et al. 2011, Corstophine et al. 2010, Carmona et al. 2013). In many taxa, these trans-Arctic affinities remain unexplored and their taxonomic status still reflects the particular scientific tradition rather than the true systematic relationships.
1.1.3. Paleogeography and climatic history

In principle, opportunities for trans-Arctic dispersal may have existed recurrently after the Miocene. They have been controlled by the hydrological and climatic events, together with the ecology of organisms with different temperature tolerances and dispersal characteristics. Up to the Late Miocene, the biotas of the Northern Pacific and Atlantic Oceans had developed independently for over 20 Myr and they did not share any common species (e.g. Dall 1920, MacNeil 1957, Marincovich 1993, Nolf and Marincovich 1994, Marincovich and Gladenkov 2001). The Bering Strait was submerged 5.5–5 Myr ago (Marincovich and Gladenkov 2001, Salzmann et al. 2008, Raymo et al 2011), allowing faunal connections between the northernmost parts of the boreal Pacific and Atlantic Oceans for the first time. The Late Miocene and Early Pliocene were characterized by high sea levels and temperatures, which could have enabled the trans-Arctic movements of the boreal fauna for a prolonged period, possibly for two millions of years. The trans-Arctic dispersal peak documented in the fossil record, the Great Trans-Arctic biotic Interchange (GTAI; Durham and MacNeil 1967, Einarsson et al. 1967, Vermeij 1989, 1991, Briggs 1995), is assumed to have taken place during the Middle Pliocene ca 3.5 Myr ago. A single more or less coincident
interchange event has been often assumed, but possibilities for exchange might have occurred also before and after the GTAI.

Towards the end of the Pliocene, the Quaternary Ice Ages began gradually (Flesche Kleiven et al. 2002), severing the Arctic connection. Three major periods with high sea levels are recorded during the early stages of Ice Age, in the deposits of Alaska and Greenland. Waters that flooded the Bering Strait putatively permitted the exchange of water masses (and animals) between the oceans. These events are dated to 2.7 Myr, 2.6 Myr and 2.4–2.2 Myr, although other authors have placed the latter highstand to 1.4 Myr instead (Miller et al. 2010 and references therein). The highstands were followed by a period of non-marine deposition within the northern Bering, Chukchi and Beaufort Sea coasts, reflecting long-term global cooling and plausibly lengthy vicariance between the boreal faunas (Kaufman and Brigham-Grette 1993). Apparently, the next time relative sea level reached above the present levels was the Middle Pleistocene interglaciation MIS11 ca 400 kyr ago (Droxler & Farrell 2000, Olson and Hearty 2009) and again during the Last Interglaciation MIS 5e ca 130–110 kyr ago (e.g. Kukla 2000, Anderson et al. 2006). Both stages were potentially favourable for trans-Arctic dispersal.

The final, most recent opening of the Bering Strait took place ca 13–11 kyr ago (Kaufman et al. 2004, Keigwin et al. 2006) and already ca 13 kyr ago the cold-tolerant Pacific molluscs *Hiatella arctica* and *Portlandia arctica* entered the Canadian Arctic (Gordillo and Aitken 2000, Kaufman et al. 2004). By 11.5 kyr ago the boreal thermophiles *Mytilus edulis* and *Macoma balthica* had spread northwards from the Bering Strait, coinciding with the entry of both the Atlantic and Pacific bowhead whales to the Canadian Arctic archipelago (past their present ranges; Kaufman et al. 2004). This marks the local Holocene Thermal Maximum (HTM; Renssen et al. 2012), which varied regionally and for instance was reached only 7.8–6.8 kyr ago in the European Arctic (Duplessy et al. 2001, 2005). Following the HTM, temperatures have decreased gradually (e.g. Miller et al. 2010), probably reducing the trans-Arctic dispersal potential of the boreal fauna.

### 1.2. Phylogeographical and systematic background

#### 1.2.1. Phylogeographical patterns in amphi-boreal taxa

It is known from the fossil record that the direction of trans-Arctic dispersal has taken place mainly from the Pacific to the Atlantic Ocean (cf. Vermeij 1991). Phylogeographical studies of the amphi-boreal taxa have so far demonstrated a variety of inter-oceanic systematic affinities among the vicariant taxa and populations. The variety of phylogenetic relationships in the amphi-boreal marine taxa has been outlined by Vermeij (1991) and Cunningham and Collins (1998). Based on the fossil data, Vermeij (1991) recognized that intra-basin systematic diversification (cladogenesis) following the trans-Arctic
invasion is a rare phenomenon, and more typically only a single derived species resulted from colonization to the Atlantic. Cunningham and Collins (1998) distinguished three main biogeographical patterns regarding the relationships of the North Pacific and North Atlantic taxa: (i) long vicariant histories of the Pacific and Atlantic sister taxa without further invasions, (ii) repeated trans-Arctic invasions and (iii) recent genetic connections between the oceans. Subsequent studies have revealed that repeated trans-Arctic invasions have sometimes led to a secondary contact of two consecutively invading Pacific lineages within the Atlantic (e.g. Väinölä 2003, Riginos and Cunningham 2005, Coulson et al. 2006, Strelkov et al. 2007, Ursvik et al. 2007). Such secondary contacts following repeated invasions may further result in hybridization between previously vicariant sister taxa. Paleontological and molecular evidence provides examples of these patterns: the gastropod *Nucella* has diverged to sister-species in respective oceans (*N. freycinetii* and *N. lapillus*; Gibbard 1991, Collins et al. 1996), the Baltic clam *Macoma balthica* has invaded Atlantic coasts several times through the Arctic route and secondary contacts of distinct invasions have taken place (Väinölä 2003, Nikula et al. 2007), whereas the green sea urchin *Strongylocentrotus droebachiensis* shows practically no differentiation between the Pacific and Atlantic Oceans (Addison and Hart 2005). Further population genetic and phylogeographical studies that examine the trans-Arctic connections in particular species or groups are now accumulating, but a comparative documentation on the genetic outcomes of trans-Arctic connections is so far missing (such as the intra-oceanic studies within Atlantic or Pacific; e.g. Wares and Cunninham 2001, Ilves et al. 2010, Keyse et al. 2014). If the patterns of repeated dispersal and secondary contacts turn out to be common in the large-scale data, our understanding of the genetic background of boreal marine fauna should be changed significantly.

1.2.2. Genetic introgression
Recent research has shown that hybridization between taxa and subsequent mixing of gene-pools (introgression) may be more common in the marine environment than previously believed (reviewed by Arnold & Fogarty 2009). Examples of introgressive hybridization following the trans-Arctic secondary contacts involve the *Macoma balthica* and *Mytilus edulis* bivalve complexes (Väinölä 2003, Riginos & Cunningham 2005, Strelkov et al. 2007) and capelin *Mallotus villosus* (Colbeck et al. 2011). Introgression might have either negative or positive impact on taxa. The potential consequences include loss of genetic identity (reverse speciation), which could be disadvantageous in populations where strong local adaptation has taken place (Seehausen 2006). It might also reduce the genomic stability of the resulting mixture and thus weaken the ecological functioning of populations (e.g. cause problems in reproduction). Alternatively, the new genetic material may offer a rich source of potential adaptive variation, including new genes from the invading taxa and novel gene combinations in the resulting mixture. Thus, the secondary contacts could also produce totally new genetic diversity via fusion of taxa (cf. Nikula et al. 2008, Frade et al. 2010, Stemshorn et al. 2011).
1.2.3. Cryptic systematic diversity in marine fauna

Molecular markers provide new and accurate taxonomic characters and have revealed a host of new, previously undetected taxonomic diversity in marine environments (e.g. Knowlton 2000, Appeltans et al. 2012). Such cryptic taxa are either morphologically unidentifiable from their sister taxa or the distinctive characters have gone unnoticed. Especially the species, for which cosmopolitan or amphi-oceanic distributions have been assumed, have almost regularly been shown to consist of complexes of several cryptic species-level taxa. Some examples include jellyfish (Dawson & Jacobs 2001), sipunculan worms (Kawauchi & Giribet 2010), marine nematodes (Derycke et al. 2005) sponges (Xavier et al. 2010) and crustaceans (Cabezas et al. 2013). The detection of such cryptic species-level units is especially notable for the biodiversity of relatively species-poor regions, e.g. the Baltic Sea or other boreal–arctic seas, which were recently disrupted by glaciations. Resolving the diversity and phylogeny in such widespread taxa also provides a basis for a broader biogeographical understanding.

1.3. Study systems

The thesis examines the amphi-boreal phylogeography at different systematic levels in two model taxa, a species pair of clupeid fishes *Clupea* spp. and the bivalve mollusc genus *Hiatella*, and in a multi-phylum comparative framework based on new data and existing literature.

1.3.1. Pacific and Atlantic herrings – an example of amphi-boreal species pair

The Pacific and Atlantic herrings (*Clupea pallasii* and *Clupea harengus*) make one of the most prominent vicariant amphi-boreal taxon pairs. These fishes are pelagic planktivores, abundant on both coasts of their respective oceans from temperate to subarctic regions. The two sister species, which are separated by the Arctic dispersal barrier, have diverged in their ecology and are characterised by substantial differences in their breeding habits. They are morphologically similar, although also some differences exist, most notably in the number of vertebrae (Svetovidov 1952). They can be accurately distinguished from each other also by genetic markers (Grant 1986, Jørstad et al. 1994, Jørstad 2004). However, the inter-oceanic vicariance is not complete, but is broken by the presence of remote populations of the Pacific herring in the NE Atlantic, particularly in the White Sea and the south-eastern Barents Sea (Svetovidov 1952, Jørstad 2004). Even further west, isolated occurrences in some Norwegian Sea fjords are known (Hognestad 1994, Jørstad et al. 1994). The distribution of the Atlantic herring extends into these same northeast European seas during the feeding migrations, while the breeding areas are thought to be separate (Svetovidov 1952, Jørstad and Pedersen 1986, Lajus et al. 2007). The European Pacific herring has been thought to comprise two subspecies (the White Sea herring *C. pallasii marisalbi* Berg, 1923 and the Chesa–Pechora herring *C. pallasii suworowi* Rabinerson, 1927; e.g. Novikov et al. 2001). Additionally, the European Pacific herring populations display

1.3.2. The near-cosmopolitan bivalve *Hiatella arctica*

Inaccurate systematic classifications can complicate biogeographical inferences at different temporal scales. In invertebrates, the cryptic diversity usually extends deeper in terms of evolutionary time than in vertebrates. As an example of resolving inter-oceanic relationships on both deeper and more recent time-scales, we examined the systematic diversity and phylogeny of the bivalve *Hiatella arctica* (sensu lato). It is one of the most widespread marine benthic invertebrates (Hayward et al. 1995, Kozloff 1996, Coan et al. 2000) and has often been characterised as a cosmopolitan species (e.g. Bernard 1979, Lubinsky 1980), occurring in the in arctic, antarctic, boreal, temperate and subtropical zones. However, *Hiatella* is most commonly found in the circum-boreal and circum-arctic regions and represents an important member of the amphi-boreal-arctic faunal element.

The systematics of *Hiatella* has long been in the state of confusion and consistent relationship between the morphology, ecology and the suggested taxonomical units is missing (Coan et al. 2000, Oberlechner 2008, Mikkelsen 2011). On one hand, the overall morphology of *Hiatella* has not essentially changed during the last 150 Myr (Kelly 1980, Gordillo 2001) and on the other hand, the external morphology is rather variable even within populations. The characters (unsuccessfully) suggested to be associated with taxonomy include spawning time, egg colour dimorphism, larval and postlarval morphology, the presence of byssus threads, the position of adductor muscles and the colour of the siphon tips (Coan et al. 2000). *Hiatella* has planktotrophic larvae and high dispersal capacity and it shows great plasticity in the habitat choice and the mode of life (Gordillo 2001). *Hiatella* can be found from the intertidal to the depths of several hundreds of meters. Adults can attach themselves with byssus threads to other invertebrates (Narchi 1973) or within the holdfasts of algae (e.g. Lippert et al. 2001), live nestled inside sponges (Micali & Solustri 2004), bored inside limestone or sandstone (Hunter 1949) or encrusted as epibionts on other invertebrates (Gallardo & Osorio 1978).
2. OBJECTIVES

The opportunities for trans-Arctic marine migration will most likely be enhanced by the current warming climate and increasing Arctic shipping. However, we are only beginning to understand the significance and dynamics of the trans-Arctic dispersal that has moulded the composition and relationships of the boreal marine biota before the human intervention. The general objective of this thesis is to shed light on the dynamics of trans-Arctic movements of marine fauna within the timescale from the Pliocene to the present and to understand the consequences of these movements on the present systematic and genetic diversity within the amphi-boreal faunal element. I have taken a molecular perspective to explore the patterns of invasions to new areas and migration among populations. Further, I will discuss the genetic outcomes and biological significance of such events.

The thesis is based on four original papers. The general objectives outlined above are broadly addressed in the final paper IV, which is a comparative phylogeographical study that gathers data on a range of systematically and ecologically different marine animals, including bivalves, gastropods, crustaceans, echinoderms, polychaetes, fishes and mammals. A comparative data set on mitochondrial lineage diversity, comprising subsets from the different phyla, is compiled in order to examine the general patterns of trans-Arctic faunal exchange over the time-scale from the Pliocene to the present. These kinds of biological records of genetic divergence can then be paralleled with geological and climatic references. More specifically, the objectives were to recognize (i) the possible multiplicity of dispersal events in single taxa (genus or species) and (ii) to estimate the temporal distribution of dispersal times. At the same time, the study aimed to document (iii) the potential unrecognised cryptic taxa that have resulted from trans-Arctic dispersal, (iv) the secondary contacts of diverged trans-Arctic lineages and (v) the evidence for cases of introgression following such contacts.

Three more specific case studies were targeted at phylogeography and systematics of two exemplary amphi-oceanic taxa, i.e. the herrings and the Hiattella bivalves, introduced above. Paper I is a phylogeographical study on an ecologically important amphi-boreal marine fish. The specific aim was to trace the biogeographical history of the Pacific herring occurring in NE Europe. Particularly, the timing and demographic history of the invasion were targeted. The genetic sub-structuring within the relatively recently invaded region and the “native” range in the Northern Pacific was studied and also compared to the resident Atlantic sister-species, the Atlantic herring. Paper II presents evidence and analyses of the genetic introgression from the Atlantic herring into the Pacific herring that has followed their secondary contact in NE Europe. The specific questions involve the age and depth of introgression (equilibrium structure of mixed populations) and the geographical variation in the degree of introgression. Paper III expands the time horizon deeper into the past, as it explores the molecular
systematics and biogeography of the nearly cosmopolitan bivalve complex *Hiatella*, whose distribution also covers the typical amphi-boreal range. The aim of this study was (i) to recognize the true systematic diversity and basic distribution of the component taxa of *Hiatella* in the Northern Hemisphere, (ii) to resolve the phylogeny of the (anticipated) newly detected cryptic species, (iii) to work out a hypothesis of their biogeographical history and (iv) to identify the subset of taxa that represent the amphi-boreal faunal element and to trace its trans-Arctic invasion history within the appropriate time scale from the Pliocene to the present.
3. MATERIAL AND METHODS

3.1. Material

A wide-ranging sampling of a selected set of amphi-boreal and boreal-arctic fish and invertebrate taxa was carried out. Most of the samples were collected by the project members during 2007–2013. Material was also obtained through collaborators, colleagues, friends and occasionally by purchase from around the world. Additionally, some material was obtained from natural history museums (see Acknowledgements). New molecular data sets were produced from altogether 22 genera, including nine fishes (Clupea, Eleginus, Gasterosteus, Liopsetta, Lumpenus, Limanda, Myoxocephalus, Platichthys and Zoarces), twelve invertebrates (bivalve molluscs Hiatella, Macoma calcarea, Modiolus modiolus, Turtonia minuta, gastropod mollusc Aeolidia papillosa, echinoderm Echinarachnius parma, priapulid worm Priapulus caudatus, crustaceans Balanus crenatus, Semibalanus balanoides, Calliopius laeviusculus, Gammarus setosus, Pandalus borealis) and one mammal, the ringed seal Phoca hispida. We aimed at sampling each of the four boreal continental margins (NW Pacific, NE Pacific, NW Atlantic, NE Atlantic), while this was not always possible. In the Pacific, sampling from the seas on both sides of the Bering Strait was considered important (North Bering, Beaufort and Chukchi Seas), since these regions represent the first step of a trans-Arctic route towards the Atlantic Ocean. Where high lineage diversity was initially discovered, denser sampling was sometimes carried out. In particular, the Skagerrak region provided such a case for Hiatella (III). Extensive sampling was also carried out for the exemplary case of the Pacific herring, both in the native range and in the NE European seas (I, II). Further, supplementary molecular data from new regions was produced to accompany data from previous phylogeographical studies of the sea urchin Strongylocentrotus droebachiensis, polynoid scale worm Harmothoe imbricate, mysid Neomysis spp. and the mussel Mytilus trossulus. Either tissue samples (pieces of fin or muscle from fish) or entire animals (invertebrates) were stored in 96% ethanol.

In addition to the new data, we compiled published sequence data on amphi-boreal or arctic-boreal taxa from public repositories (GenBank, BOLD), based on existing literature. In order to enable comparisons among taxa, only studies with comparable mitochondrial sequence data were included (IV).

3.2. Molecular markers

This thesis is based on information provided by the variation in molecular markers. Markers used for a particular study should be selected based on the study questions and evolutionary time scales concerned. Mitochondrial DNA (mtDNA) is still today the most useful tool available for comparative phylogeographical studies (cf. Rubinoff
and Holland 2005, Avise 2009) which deal with events on relatively recent evolutionary time scales and low taxonomic levels (within species and genera) (e.g. Bouzat 2000). Protein coding mtDNA genes can be assumed to show a relatively steady accumulation of mutations, at a comparable rate within a given group of organisms (within phyla or classes). Furthermore, the high mutation rate and variability in mtDNA makes it very informative in studies of intra-specific populations or closely related species (e.g. Avise 2009). However, mtDNA studies involve some well-recognized caveats (e.g. Bazin et al. 2006, Dowling et al. 2008, White et al. 2008): mtDNA only reflects the history of maternal lineages and it is particularly liable to inter-specific introgression that can confound the correspondence of mtDNA and taxonomic identities (relevant amphiboreal examples in Strelkov et al. 2007, Väinölä & Strelkov 2011). Furthermore, as mtDNA is a single nonrecombinating molecule, selection on any gene and particularly on a new advantageous mutation will affect the frequencies of haplotypes of other mtDNA genes too, whence the expectations of neutral evolution will be disturbed (e.g. Gillespie 2000, Galtier et al. 2009). In a selective sweep, a positively selected mutation will soon draw into fixation its mitochondrial haplotype and wipe out variation at all the other loci (and nucleotide sites). This would compromise any interpretations of demographic history, and of short-term haplotype replacement rate. Nevertheless, the amount of neutral substitutions accumulating in vicarious lineages should still reflect the time elapsed over a longer period. A further confounding factor may be the presence of two independently inherited mitochondrial lineages as documented in several bivalve molluscs (doubly uniparental inheritance; Skibinski et al. 1994, Passamonti & Scali 2001) or the more general presence of nuclear pseudogenes (Leister 2005). All these pitfalls need to be taken into account when evaluating the implications of the results. In order to ultimately resolve the phylogeny and the breeding population structure of any taxon, data on other genetic markers are also needed.

In the herring study, we examined mtDNA sequences of the nearly entire cytochrome b gene (cyt-b, 1131 bp) and part of the control region (CR, 481 bp) (I, II) and a set of nuclear allozyme markers for a part of the material (II). Allozymes have so far provided the most powerful nuclear markers for distinguishing between these two sister species (e.g. Grant 1986, Jørstad et al. 1994) and they were used for species identification and for a study of potential interbreeding in areas of range overlap. In Hiatella (III), both mitochondrial and nuclear markers were studied, in order to resolve also deeper (older) nodes in the phylogeny and to gain information on the true systematic structure of the species complex (not just matrilines). In addition to the mitochondrial cytochrome c oxidase subunit I (COI, 653 bp) we studied the nuclear adenine nucleotide translocase (ANT, 335 bp) and the nuclear 28S rDNA that codes for the large subunit ribosomal RNA (28S, 678 bp). For the comparative study across taxa (IV), sequences of particular mitochondrial genes were used to provide comparable data across a wide set of taxa. These genes were COI for invertebrates (the widely used barcoding marker) and cyt-b for fishes and mammals, the ones most widely available also in the reference literature. Additionally, the available COI data of fishes was reviewed.
3.3. Data analyses

3.3.1. Molecular diversity and phylogenetic analyses

The diversity and genealogical relationships of mitochondrial haplotypes of the Pacific and Atlantic herrings were illustrated with neighbor-joining trees from pairwise distances between haplotypes estimated in PAUP* 4.0 (Swofford 2003) under the best-fit substitution models suggested by modeltest (JModelTest 2.1.4, Darriba et al. 2011) (I). The diversity and relationships of haplotype lineages of the Hiattella complex, representing a wider evolutionary divergence (III), were illustrated with Bayesian inference consensus trees, reconstructed separately for COI, ANT and for a concatenated dataset of all three genes, using MrBayes 3.1 (Ronquist & Huelsenbeck 2003) and substitution models selected. The 28S data, which involved little variation, were illustrated with a maximum parsimony tree. In the comparative study of the trans-Arctic differentiation (IV), the genealogical relationships of mtDNA haplotypes were illustrated with neighbor-joining trees constructed in MEGA 6 (Tamura et al. 2013). To allow more straightforward comparisons among taxa, a single nucleotide substitution model was applied for all of the taxa. This was selected by screening alternative models in modeltest for several of the datasets. The one that was the best fitting or among the best ones in the majority of cases was chosen.

The standard intra-population molecular diversity statistics for the mtDNA sequence data were calculated using DnaSP v5 (I, II, III). Genetic differences among geographical regions and among populations within regions were assessed using the $\Phi_{ST}$ statistics of the analysis of molecular variance (AMOVA) using the Arlequin software (I, III, IV; Excoffier & Lischer 2010). Inter-population relationships of European Pacific herring were illustrated by metric multidimensional scaling (MDS) from the pairwise $\Phi_{ST}$ distance matrix, using NTSYS-pc software (Rohlf 1990). The demographic histories (changes in effective population size) in various herring populations were traced from the mismatch distributions of the pairwise sequence differences and further by coalescence simulations using Bayesian skyline plot analysis (BSL) with BEAST v1.5.3 software (Drummond et al. 2012) (I).

3.3.2. Intraspecific patterns and recent trans-Arctic connections

When intraspecific mitochondrial lineages included individuals from both the Pacific and Atlantic basins, that subset of data (i.e. the shared lineage) was subjected to further analyses. A statistical parsimony haplotype network was constructed using the TCS 1.21 software, in order to illustrate the distribution of the diversity in a phylogenetic context (I, III, IV; Clement et al. 2000). Where sample sizes were sufficient, the most recent trans-Arctic invasion event and levels of gene flow after that were estimated with a MCMC coalescence model simulation implemented in the program IM (I, III) or IMa (IV) (Nielsen & Wakeley 2001, Hey and Nielsen 2007). These simulations assume a population model where an ancestral population splits into two descendant
populations, which subsequently diverge but can still be connected by gene flow. The molecular distances representing the split were estimated also as net distances (to account for the ancestral polymorphism approximated by the current intra-population variation) using MEGA (IV).

### 3.3.3. Substitution rates and divergence time

Estimates of realized nucleotide divergence between lineages were used as surrogates of relative time since the split of populations, which was most often equated with the trans-Arctic dispersal events in the focus of the study. To strive for a linear relationship, appropriate substitution models (chosen using the modeltest program, Darriba et al. 2011) were applied.

To relate the estimated divergence to absolute and geological time, estimates of the actual nucleotide substitution rate are needed. Although such estimates are at the best very coarse (most likely inaccurate), they are useful at least in providing a basis for discussion of the data in a context comprehensible in the framework of geological history. The estimates used were based on: (i) rates presented in literature for comparable data sets (I), (ii) calibration by matching the (corrected) molecular divergence in chosen taxa with the Great Trans-Arctic Interchange that followed the Pliocene opening of the Bering Strait, which has generally been dated to 3.5 Myr and often used as a basis of calibration in trans-Arctic phylogeographical studies (I, III, IV) and (iii) connecting the appropriate observed lineages to the first documented fossil occurrences of the species (complex) in the corresponding ocean basins (III). From previous and from the present data, the general rates of mitochondrial divergence are inferred to vary considerably among major taxonomic groups, but are assumed to be relatively similar within the groups. Group-specific rates were therefore applied in the multi-taxon comparisons.

None of the approaches above comes without problems. Biogeographical events used for calibration are usually simplified to a certain date, although they in reality would represent longer time periods. As regards the Pliocene GTAI, a plausible window for exchange could range from 5.5 to 2.4 Myr ago, instead of the often cited 3.5 Myr; the predominance of the conventional date has been argued based on the change in oceanic circulation patterns in the Bering Strait region (Haug and Tiedemann 1998). Divergence rates in the literature are rarely based on calibrations on fossils, which could be solidly linked to genetic data (i.e. the fossils used for calibration might not represent the extant lineage under consideration), which is also the case with our *Hiatella* data (III). Moreover, a phenomenon known as the *time dependency of molecular rates* causes additional problems if dealing with different time-scales (review in Ho et al. 2011). This phenomenon, inferred in many studies from a variety of taxa, is seen as an apparent acceleration of substitution rates when dealing with more recent time scales vs. those representing millions of years, in spite of model corrections which ought to linearize the divergence vs. time relationship. For instance, in marine invertebrates
Crandall et al. (2012) obtained three- to six-fold higher substitution rates when using calibration dates younger than 5 Myr vs. those based on older dates.

3.3.4. Taxon delimitation
The data from the *Hiatella* bivalve complex (III) disclosed several lineages, both allopatric and sympatric, that were separated by substantial nucleotide distances, suggesting a presence of distinct taxonomic subdivisions. In addition to verifying character congruence (whether the mtDNA and nuclear lineages matched each other at the individual level), we applied two formal approaches to genetic species limitation, based on the coalescence. GMYC (referring to the General Mixed Yule Coalescent model) is a maximum likelihood approach to discriminate population and speciation patterns from mitochondrial genealogies alone (Pons et al. 2006), whereas Bayesian Species Delimitation was conducted on the three-marker dataset (28S, ANT, COI) using the program BPP v2.2 (Rannala & Yang 2003, Yang & Rannala 2010).

3.3.5. Coincidence of phylogeographical split times
We used hierarchical approximate Bayesian computation HABC (Beaumont et al. 2002, Hickerson & Meyer 2008) as implemented in mBayes (Hickerson et al. 2007, Huang et al. 2011) to test the hypotheses of simultaneous divergence vs. multiple divergence times among the various vicarious pairs of taxa with amphi-boreal distributions (IV). Only fishes were included in the analysis, to keep the analysis simple and to minimize the complications related to taxon specific mutation dynamics. *Cyt-b* datasets of altogether 16 pairs of boreal or arctic-boreal fish taxa were used for the analysis.

3.3.6. Analyses of introgression
The mixing and potential introgression between the Atlantic and Pacific herrings (II) in North European populations was assessed at the population level first by simple estimates of the average proportions of species-specific alleles, using frequencies of “known” taxonomically pure samples as reference. The admixture of the parental genomes in individuals was analysed using several approaches, including Bayesian assignment of individual genomic ancestries (STRUCTURE 2.2; Pritchard et al. 2000), conventional statistics of intergenic equilibria/disequilibria within samples (Asmussen & Basten 1994, Lewis & Zaykin 2002) and comparing the observed distribution of individual multilocus genotypes to the distributions expected in panmictic population and in non-hybridizing mixture of two species (e.g. Väinölä & Strelkov 2011). The previously published allozyme frequency data of North European herrings from altogether 53 population samples, comprising material from both species, reported in Jørstad et al. (1991, 1994), Jørstad (2004) and Semenova et al. (2009), were re-analysed by principal component analysis (PCA) in PAST (Hammer et al. 2001). The main axis (PC1), which in effect represented a weighed sum of the typical parental allele frequencies in each sample, was scaled to give estimates of the population-level admixture or introgression.
4. RESULTS AND DISCUSSION

4.1. The diversity of inter-oceanic genetic distances

From the new multi-taxon dataset and the compilation of the previously published mitochondrial data, encompassing comparisons in altogether 70 amphi-oceanic species or species groups, a wide distribution of Pacific–Atlantic divergence estimates was documented. This suggests that trans-Arctic faunal dispersal has taken place at multiple time frames and in several cases repeatedly after the initial opening of the Bering Strait (IV). Within each of the major taxonomic groups examined (fishes, molluscs, crustaceans, echinoderms, polychaetes), the distribution of divergence estimates was broad, involving up to 10–20 fold differences among taxa. Similarly divergence estimates varied considerably between taxa representing groups of relatively closely related and presumably ecologically similar taxa. Successive divergence are inferred to have taken place at different times also within a single nominal species (e.g. the *Macoma balthica* complex; Nikula et al. 2007). For instance, in flatfishes of the family Pleuronectidae, the range of Pacific vs. Atlantic cyt-b divergence estimates included 4.8% between sister the species pair *Platichthys stellatus* / *P. flesus*, 4% between *Hippoglossoides robustus* / *H. platessoides*, 3.13% between *Limanda aspera* and *L. limanda*, 1.98% between *Hippoglossus stenolepis* and *H. hippoglossus* and incomplete lineage sorting (~ 0%) between Pacific and Atlantic basin populations of *Liopsetta glacialis*. Plotted against a putative evolutionary time-scale, the model-corrected inter-oceanic net distances represent divergences ranging from the Late Miocene to the Holocene (Fig. 2; taxon specific divergence estimates in Table 3 of IV). In fishes, the majority of estimates would well fit the mid-Pliocene time frame of the GTAI, but several substantially younger and older estimates were also recorded. The spread of distance estimates in e.g. mollusc taxa was even more prominent, whereas for instance the crustacean estimates tended to be in the lower (Pleistocene) end of the scale.

The variable Pacific–Atlantic mtDNA divergence estimates at the deep end of the scale, together with the varying dates of first appearance of Pacific taxa in the Atlantic fossil record, point to extended dispersal possibilities through the Arctic basin during the Pliocene. For example, the COI divergence (38.7% TN+Γ net distance) and first Atlantic fossils (Early Pliocene 5.3–3.5 Myr; Kafanov 1987, Vermeij 1991) of the boreal bivalve *Mytilus edulis* sensu lato consistently point to an older invasion than the COI divergence (22%; Nikula et al. 2007) and the fossil reference (2.4 Myr; Simonarson et al. 1998, Funder et al. 2001) of another boreal bivalve *Macoma balthica*. Apparently, the initial establishment of the trans-Arctic connection in the Pliocene was followed by a prolonged period of invasions or even continuous gene flow. Many boreal taxa plausibly extended their distributions to the Arctic basin, and the geographical distance between Pacific and Atlantic populations was shorter than today.
Results and Discussion

Figure 2 Trans-Arctic net divergence estimates (from TN+Γ substitution model, where 𝛼 = 0.2; from mitochondrial COI data for invertebrates, cyt-b for fishes), based on the tentative group-specific molecular divergence rates: 11.1% Myr⁻¹ (for bivalves; yellow crosses), 6.1% Myr⁻¹ (for gastropods; blue crosses), 6.0% Myr⁻¹ (for polychaetes and other worms; red crosses), 4.4% Myr⁻¹ (for echinoderms; green crosses), 4.3% Myr⁻¹ (for crustaceans; purple crosses) and 1.1% Myr⁻¹ (for fishes) (see IV for details). Examples of different phylogeographical patterns from three bivalves are illustrated with neighbor-joining trees (see below). Globally distributed benthic δ¹⁸O records (Lisiecki and Raymo 2005), indicating the pattern of climate variation over the past 5 million years, are plotted on the same scale.
The onset of the Pleistocene Ice Ages closed the connection between boreal faunas for most of the time, but highly variable Pacific–Atlantic divergence estimates imply that substantial amount of trans-Arctic dispersal has taken place also during the marine transgressions of Pleistocene interglacials. The HABC approach suggested seven different divergence dates for the fish dataset alone (IV). Some potential peaks in the distribution of estimates may be seen (e.g. for invertebrates around 0.3 and 1.5 Myr on the scale given), which might be associated with particular Pleistocene interglacials. It must be noted that in the framework of intra-generic comparisons, inferences of multiple invasions are not dependent on any divergence rate calibrations. Also, due to the mostly sparse sampling in the dataset, cases of repeated dispersal could easily have gone unnoticed and the frequency of taxa with multiple invasions should be an underestimate.

4.2. Phylogeographical patterns in amphi-boreal taxa

The inferred phylogeographical patterns were classified into four main categories judged to represent different inter-oceanic isolation-dispersal histories (a–d in Table 1; taxon specific pattern designations in Table 1 of IV). Among the 70 amphi-boreal or arctic-boreal taxa that were examined in the synthetic comparative study (IV), we recorded altogether 34 taxa with simple dichotomous vicariant histories between the oceans (patterns a and a1; e.g. the bivalve Modiolus modiolus; Fig. 2) and 19 taxa with a minimum of two distinct trans-Arctic divergence and inferred dispersal events (pattern b; e.g. the Macoma balthica bivalve complex; Fig. 2). In the genera with repeated invasions, in all except one a secondary contact of two succesively invading lineages was observed, usually within the Atlantic basin (pattern c; e.g. M. balthica, Fig. 2). Altogether 28 taxa showed incomplete lineage sorting between the

Table 1. Frequencies of different phylogeographical patterns in amphi-boreal(-arctic) taxa in different animal groups, in the compilation of mitochondrial sequence data (IV).

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Molluscs</th>
<th>Crustaceans</th>
<th>Echinoderms</th>
<th>Worms</th>
<th>Fishes</th>
<th>Seals</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>14</td>
<td>25</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>a1</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>b</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>c</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>13</td>
<td>11</td>
<td>19</td>
<td>40</td>
<td>3</td>
<td>104</td>
</tr>
</tbody>
</table>

a - long vicariant histories (descending from the Pliocene–Early Pleistocene)
a1 - more recent vicariant histories (descending from the Middle Pleistocene, 1–0.2 Myr ago)
b - minimum of two distinct trans-Arctic divergence events
c - inferred secondary contact following repeated trans-Arctic dispersal
d - incomplete lineage sorting or close relationship, interpreted as late- or post-glacial dispersal
e - hypothesized post-glacial contact within Arctic Ocean, following short vicariance (e.g. in fishes since 100 ky ago)
Results and Discussion

23

oceans, suggesting recent dispersal or efficient ongoing migration (pattern d; e.g. the chalky macoma *M. calcarea*, Fig. 2). Yet another pattern was recorded in some arctic-boreal fishes, where vicariant Pacific and Atlantic lineages have re-entered the Arctic in post-glacial times and meet there in secondary contact (cf. Mecklenburg et al. 2011, IV).

4.2.1. Vicariant histories without further dispersal

An early vicariance event (plausibly Pliocene/Early Pleistocene) without evidence of later contact was inferred (pattern a) for example in the bivalves *Modiolus modiolus* and *Turtonia minuta*, clades in the gastropod genera *Littorina* (Reid et al. 1990, Wares & Cunninham 2001, Reid et al. 2012, Layton et al. 2014) and *Nucella* (Collins et al. 1996, Wares & Cunninham 2001, Marko et al. 2014) and the sea slug *Alderia modesta*. Among crustaceans, vicariance is inferred for the barnacle *Semibalanus balanoides*, the amphipod *Calliopius laeviusculus* and mysids of the genus *Neomysis*. The pattern was also recorded in *Asterias* sea stars (Wares 2001, Matsubara et al. 2005), as well as several polychaetes including the boreal *Hediste* spp., *Alitta virens* and several others from the Canadian study of Carr et al. (2011). In fishes, similar instances of deep vicariance include inter-oceanic sister taxa of the gadid genus *Eleginus*, the pholid genus *Pholis* (Steinke et al. 2009), the flatfishes *Hippoglossus*, *Hippoglossoides*, *Limanda* and *Platichthys* (e.g. Mecklenburg et al. 2011) as well as cottoid clades in the sculpin genera *Gymnocanthus* and *Myoxocephalus*. Dichotomies representing similar but plausibly more shallow vicariant events (pattern a1), would in turn be indicated for several sea star genera (Corstrophine et al. 2010), for zoarcid fishes (*Zoarces* eelpouts) and for the largely freshwater stickleback fish genera *Gasterosteus* and *Pungitius*. However, the latter two may involve several equally deep clades within each basin and the frequency of exchange events is difficult to resolve.

4.2.2. Several trans-Arctic divergence events

The 19 instances where multiple trans-Arctic invasions within a taxon were inferred (pattern b) included the bivalves *Macoma balthica* (Nikula et al. 2007), *Mytilus trossulus* and *Hiatella* spp. (III), the sea slug *Aeolidea papillosa*, the crustaceans *Gammarus setosus* and *Balanus crenatus*, polychaetes *Harmothoe imbricata* (Hardy et al. 2011), *Nereis pelagica* (Carr et al. 2011) and the priapulid worm *Priapulus caudatus*. Further, this pattern involves the fishes *Aspidophoroides monopterygius* (Mecklenburg et al. 2011), *Clupea pallasii* (II), *Gadus macrocephalus* (= *G. ogac*; Coulson et al. 2006,) and *Gadus chalcogrammus* (= *G. finnarchicus*) (Ursvik et al. 2007), *Lumpenus fabricii* (Mecklenburg et al. 2011, this study), *Mallotus villosus* (Dodson et al. 2007), *Myoxocephalus scorpius* (Mecklenburg et al. 2011, this study) and the harbor and spotted seals *Phoca* spp. Instances where both of the successive invasions were inferred to have taken place on a relatively recent time frame (during the previous and current interglacials), included the echinoderms *Echinarchnium parma* and *Strongylocentrotus*
droebachiensis (as already suspected by Addison and Hart 2005), the bivalve *Hiatella* sp. L (III) and the barnacle *Balanus crenatus* (in addition to an inferred human-mediated invasion).

### 4.3. Cryptic amphi-boreal taxa

While amphi-oceanic lineages judged to be of Pliocene age usually correspond to well-established morphological sister taxa in fishes, the invertebrate lineages of similar age have most often not been recognised in taxonomy. Our multi-taxon dataset (IV) involved several instances of such cryptic lineages, newly detected or previously known: in the bivalves *Turtonia minuta* (IV) and *Modiolus modiolus* (Halanych et al. 2013, IV), in the nudibranchs *Aeolidia papillosa* (Carmona et al. 2013, IV) and *Alderia modesta* (Ellingson & Krug 2006), in the amphipod *Calliopius laevisculus* (IV), in the barnacle *Semibalanus balanoides* (Zullo & Marincovich 1990, IV) and in several polychaetes (Carr et al. 2011). The taxonomic status of these lineages requires further investigation of morphological and nuclear characters, but will probably in many cases lead to formal recognition of separate sister taxa, as has previously been the case for cryptic taxa in *Mytilus* (McDonald & Koehn 1988) and *Macoma* (Vainölä 2003). Such unrecognized level of diversity seems to be particularly common in marine invertebrates (cf. Dawson & Jacobs 2001, Kawauchi & Giribet 2010, Xavier et al. 2010, Cabezas et al. 2013). A particular search for cryptic diversity was made in the *Hiatella* bivalve complex (Box 1; III). A rich diversity of lineages judged to represent 13 separate species was found, and from these two species, *Hiatella* L (originally North Pacific) and *Hiatella* K (Northeast Atlantic), were recognized to represent an amphi-boreal taxon pair.

### 4.4. Post-glacial dispersal

The compiled comparative data involves new evidence of post-glacial trans-Arctic connections for several taxa. These include the Pacific herring (I), the arctic flounder *Liopsetta glacialis* (IV), the cryptic bivalve species *Hiatella* L (III), the chalky macoma *Macoma calcarea* (IV), the amphipod *Gammarus setosus* (IV), the priapulid worm *Priapulus caudatus* (IV) and the ringed seal *Phoca hispida* in addition to 21 cases previously reported (see Table 4 in IV). Regarding the broad presence of this pattern among the studied taxa, the Holocene is inferred to have been an important period for trans-Arctic dispersal. These connections are thought to have preceded any human intervention, and in addition we separated cases of close relationship that were judged to represent recent human mediated introductions (IV).

In principle, the signatures obtained from mitochondrial DNA regarding recent demographic events, including both migration and population size changes, could be biased by selection on certain mtDNA haplotypes that simultaneously
Results and Discussion

Box 1. Cryptic diversity in the *Hiatella* bivalve complex (III)

Altogether 11 distinct mitochondrial lineages, separated from each other by >1% observed divergence in the COI gene (27.7–110.0% model-corrected distance; Fig. 3) and a number of lower-level subdivisions, were recognized among the 71 geographical samples of *Hiatella* from the Northern Hemisphere. The two Southern Hemisphere samples added two lineages more. Several of them were found occasionally in sympatry with others, particularly at sites in northern Europe. In the nuclear ANT gene sequence data, those same 13 lineages were consistently resolved; the ANT lineages were distinguished by 3.6–14.6% p-distance (4.2–25.9% corrected distance), whereas the 28S gene showed less informative variation. The complete mitochondrial-nuclear character congruence at the individual level, also when the lineages were found in sympatry, may be taken as evidence of taxonomic isolation in the sense of the biological species concept. Two genealogy-based approaches to species delimitation supported this conclusion. The Bayesian BPP method and the GMYC approach both recognized the same 13 taxa. With this evidence, the taxa were judged to represent putative species, tentatively designated *Hiatella* A–L.

Five of the cryptic *Hiatella* taxa were present in the European samples. At six European locations, more than one *Hiatella* species were encountered in sympatry. Presumably at least part of these co-occurrences represent very recent mixing of taxa due to human activity, as the sympatric sites mostly were harbours. *Hiatella* K was the most widespread of the European species and had the most northerly general distribution, accounting for all the arctic and subarctic sectors of the sampled European range. In contrast, we only found one species, *Hiatella* L on North American Atlantic coast (the data of Layton et al. 2014 in the BOLD database also suggest a local occurrence of *Hiatella* K in Canada). Six species were recognized from the North Pacific and the bordering sector of the Arctic Ocean. The most widespread and the most northern of these was *Hiatella* L, with a vast distribution covering the Gulf of Alaska as well as the Russian Pacific coast and extending through the Bering Sea to the Arctic Chukchi and Beaufort Seas and found again in the Northwest Atlantic.

Figure 3 The distribution of eleven cryptic Northern Hemisphere species *Hiatella* C–L and the inferred phylogeny of the *Hiatella* complex. Two Southern Hemisphere samples referred to *Hiatella* A (Strait of Magellan) and *Hiatella* B (New Zealand) are not shown. A Bayesian consensus tree based on three-gene concatenated data is shown. Branches with a posterior probability support >0.95 are indicated with thick lines. *Hiatella* L (red) and *Hiatella* K (blue) represent an amphi-boreal species pair inferred to have diverged in the Pliocene, following the opening of the Bering Strait (Fyles et al. 1991; Vermeij 1991; Gordillo 2001). Thus, the Atlantic *Hiatella* K represents a cryptic sister species to the principally Pacific *Hiatella* L, evolved in vicariance since a Pliocene trans-Arctic invasion event. Additionally, the modern distribution of *Hiatella* L shows that even more recent trans-Arctic invasions have taken place on the intra-species level.
Box 2. Post-glacial invasion and secondary contacts in herrings and in the Hiatella bivalve complex (I)

In the amphi-oceanic study of the herring mtDNA diversity, the Northeast European *Clupea pallasii* populations were inferred to represent post-glacial colonizers from the Asian–Beringian clade of original Pacific *C. pallasii* (Fig. 4a). The pattern of haplotype diversity within the European populations suggests that dispersal took place in the early post-glacial time. While the most common haplotypes are still shared with the Pacific ancestor population (incomplete lineage sorting), several new satellite haplotypes are present which seem to represent new mutations within the European populations since the colonization event.

From within the diversity of cryptic taxa in the circum-global bivalves *Hiatella* spp. (III), one species with a wide amphi-oceanic northern distribution was recognized, i.e. *Hiatella* L which is present in the North Pacific, Arctic and Northwest Atlantic Oceans. Its mtDNA haplotype network involves two distinct haplogroups, L1 and L2, of which only L1 is present in the Pacific and the adjacent Arctic, but which both co-occur in the Atlantic (Fig. 4). The average nucleotide divergence between L1 and L2 (1.2%), IM coalescence analyses and the haplotype network suggest that the co-occurrence of L1 and L2 in the NW Atlantic derives from two separate trans-Arctic invasion waves from the Pacific, plausibly during the last major interglacial (ca 110–130 kyr ago; Kukla 2000) and in the post-glacial time.

Both *Hiatella* and the herrings represent examples of taxa where repeated trans-Arctic dispersal events have taken place (pattern b) and where the latest connection was a post-glacial invasion that has led to a secondary contact with the earlier invader or resident sister taxon (pattern c). In herring, the secondary contact has led to genetic introgression from the Atlantic sister species *C. harengus*, which represents a much earlier vicariance and has evolved into a distinct species. The contact in *Hiatella* L in turn is between quite recent mtDNA lineages, for which the taxonomic identity has not been tested by nuclear genotypic data.

![Figure 4](image)

**Figure 4 (a)** Mitochondrial diversity in the sister species Pacific herring (*Clupea pallasii*) and Atlantic herring (*C. harengus*). A neighbor-joining tree from GTR+I+Γ distances of cyt-b haplotype data (1131 bp), and a statistical parsimony network (99%) of the cyt-b haplotypes in the trans-Arctic *C. pallasii* clade. **(b)** Mitochondrial diversity in the *Hiatella L – Hiatella K* species pair. A Bayesian consensus tree (GTR+I+Γ model) of haplotypes of the partial *COI* gene (653 bp), where branches with posterior probability support >0.95 are indicated with thick lines. Statistical parsimony network (99%) of the *COI* haplotypes in the amphi-oceanic taxon *Hiatella L*. The L1 and L2 subclusters of the Atlantic individuals (blue) are hypothesized to represent successive pre- and post-glacial invasions from the Pacific to the NW Atlantic. Note that the postulated rate of divergence in *Hiatella* and other bivalves is several-fold compared to that in fishes.
will affect the variation of the entire mtDNA genome (Gillespie 2000, Ballard et al. 2007). While a close mitochondrial relationship at all events will reflect recent gene exchange, in some cases it could be nonrepresentative of the whole species genome. If an invading mtDNA lineage would have an advantage over that in a resident species, it could introgress to that species and potentially spread without corresponding invasion of the nuclear genome. The study of nuclear markers too is therefore important for final evaluation of the patterns. Indeed, in many of the model cases of close interoceanic relationship and repeated invasion, congruent nuclear data do exist, often since times before mtDNA studies. These involve data from Clupea (Jørstad et al. 1994), Mytilus (McDonald & Koehn 1988, Väinölä & Strelkov 2011), Macoma balthica (Väinölä 2003) as well as Hiattella (III), Strongylocentrotus (Addison & Hart 2004) and Gadus (Árnason & Halldórsdóttir 2015). On the other hand, mitochondrial variation can be affected by new advantageous mutations and selective sweeps irrespective of hybridization and produce false signatures of population bottlenecks; this is possible in the case of invading lineages, but as well for any resident populations. Strong bottleneck effects were inferred for the invading Pacific herring in Europe (see section 4.6). In this case, introgression was inferred to have occurred from the resident Atlantic herring into the invading species at similar rate both in the nuclear and mitochondrial genomes. This could rather be seen as evidence against any strong selective sweeps that would have advanced the spread a particular Pacific haplotype.

4.5. Secondary contacts of sister taxa following re-invasion from the Pacific

Amongst the cases where more than one trans-Arctic connection was recorded (pattern b), (partly) overlapping distributions among lineages were detected in all cases but one, although sample sizes and geographical coverage in many taxa were small. The secondary arrivals of Pacific taxa to the Atlantic realm seem to have remarkably often resulted in contacts between the sister taxa or distinct lineages (pattern c; IV). Pairs of fish sister species where the current Pacific taxon now co-occurs with an Atlantic one within the NE Atlantic include the herrings Clupea pallasii/C. harengus, the cods Gadus chalcogrammus/G. morhua, the eelblennies Lumpenus fabricii/L. lampretaeformis and the smelts Osmerus dentex/O, eperlanus, all recognized also morphologically. In invertebrates, similar presence of vicariously evolved sister taxa within the NE Atlantic region have been recognized in the bivalves Mytilus trossulus/M. edulis and Macoma balthica balthica / M. b. rubra, and in unnamed cryptic lineages in the nudibranch mollusc Aeolidia papillosa, the barnacle Balanus crenatus, the amphipod Gammarus setosus and the polychaete worm Harmothoe imbricata. Similarly, pairs of sister species that both occur in the NW Atlantic sector include the fishes Aspidophoroides monopterygius/A. olrikii, Lumpenus fabricii/L. lampretaeformis and Gadus macrocephalus/G. morhua, the bivalves Mytilus trossulus/M. edulis, Macoma
Box 3. Introgression between the Atlantic and Pacific herrings

Within the Pacific herring populations of the White Sea, about 7% of the individuals were found to carry Atlantic herring mitochondria, although Atlantic herring itself is not known to breed in this basin (II). Within the local Balsfjord herring stock of the Norwegian Sea the percentage of Atlantic mitochondria was 21%, whereas no introgressed haplotypes were recorded in Pechora Sea samples. In each of the populations, corresponding levels of genomic admixture were also estimated from four diagnostic or nearly diagnostic nuclear allozyme loci. The Atlantic herring mtDNA in the White Sea occurred in individuals with practically “pure” Pacific herring nuclear background. In Balsfjord also the nuclear genomes were of a mixed composition, with similar proportions of ancestry as in the mitochondria (20%). The Balsfjord population showed no inter-locus genotypic equilibria and the introgressed mitochondria were randomly distributed among the nuclear genotypes. From this equilibrium structure and from the distribution and diversity of the introgressed haplotypes among the very diverse “native” Atlantic herring haplotypes, it was concluded that the introgression is a result of recurrent backcrossing and hybridization over a long period within the post-glacial time frame. The Balsfjord herring plausibly represents a local hybrid swarm, capable of persisting and maintaining its genetic composition without further input from the parental taxa. From a reassessment of previously published allozyme data sets, it was hypothesized that the patterns of intra-species geographical variation previously recorded in the two North European herring species may largely reflect varying levels of inter-species introgression.

Figure 5 Analysis of introgression in Northeast European herring. (a) In the genetically mixed Balsfjord herring population, the observed distribution of multilocus genotypic classes (defined by the number of Clupea pallasii-specific alleles across four allozyme loci (thick bars) is compared with the distributions expected in a panmictic equilibrium population (polygon) and in a non-hybridizing mixture of the two species (thin bars), with the same average allele frequencies. The proportion of C. harengus mitochondria in each genotype class is indicated as a black section of the frequency bar (b) A statistical parsimony network (99%) of cyt-b haplotypes in the C. harengus mitochondrial lineage. The picture shows the high diversity of the “native” Atlantic herring (black circles) and the position of the haplotypes found introgressed in C. pallasii background embedded in this diversity, separating those from the White Sea (white) and in the Balsfjord hybrid swarm (grey circles). The size of the circle or sector is proportional to the observed frequency of the haplotype.
balthica/M. petalum and further distinct (cryptic) intra-specific lineages of the bivalve Hiatella L (III), the crustaceans Gammarus setosus and Balanus crenatus, polychaete worms Eunoe nodosa, Harmothoe imbricata, H. rarispina and Nereis pelagica and of the fish Mallotus villosus.

4.6. Introgression following secondary contacts

A hypothesis of introgressive hybridization between North Pacific and North Atlantic sister taxa, resulting from secondary contacts within the Atlantic basin, was suggested by the phylogeographical analysis of the European Pacific herring mtDNA data (I). The hypothesis was then subjected to a broader analysis involving also nuclear markers (II; Box 3). Despite the long-term scientific interest in the systematic, genetic and biological variation in the European herrings (e.g. Lajus 2002, Jørstad 2004, Semenova et al. 2009), the concept of inter-species gene flow, as a factor contributing to this variation, had not been taken up previously. Introgression has been found to have considerably shaped the genetic diversity of North Atlantic populations of the boreal bivalves Macoma and Mytilus (e.g. Rawson & Hilbish 1998, Quesada et al. 1999, Riginos & Cunninham 2005, Strelkov et al. 2007). Regarding the high number of secondary contacts recorded so far in the amphi-boreal context (IV), further cases of introgression could be expected to be found in closer analyses. Overall, introgression appears to be an underrated evolutionary force in molding the diversity of the marine biota in the post-glacial environments.

4.7. Factors facilitating trans-Arctic invasion and colonization of new areas

The diversity of the divergence patterns indicates a wide variety of dispersal histories amongst members of the amphi-boreal faunal assemblage. What are the mechanisms that have enabled or restricted the dispersal and settlement of the Pacific invaders? Although generalizations concerning such a wide-ranging phenomenon may be premature, a few notions arise from the data.

According to the “ecological opportunities” concept (success of invasion is dependent on available niches; Vermeij 1991, Cunningham and Collins 1998), it has been suggested that the settlement possibilities of the Pacific taxa should differ between the western and eastern coasts of the North Atlantic Ocean (Jenkins et al. 2008). According to this hypothesis, extirpation of previous inhabitants caused by glaciations would have facilitated settlement of new trans-Arctic invaders, unless suitable Atlantic refugia for resident taxa existed near the target area. The benthic substrates are different between the coasts: northern refugia for hard-bottom obligates are assumed to have existed mainly in Europe, while soft bottom refugia have been available on both coasts, although mostly along the American coast south of Maine. However, several examples
in contrast with such invasion pattern are found in our data: the hard-bottom species *Strongylocentrotus droebachiensis* and *Hiatella* L seem to have persisted the last glaciation along the NW Atlantic coasts (although even older diversity is found in the NE Atlantic in *S. droebachiensis*; Addison & Hart 2005; III). Also, the hard-bottom bivalve *Mytilus trossulus* and the soft bottom deepwater bivalve *Macoma calcarea* have invaded both North Atlantic coasts post-glacially. Similarly in polychaetes, there are examples of both hard and soft substrate taxa with evidence of recent invasion to NW Atlantic coasts and others that lack it. Thus, this explanation based on suitable refugial environments alone is not sufficient to account for the post-glacial invasion patterns.

Regarding the post-glacial invasions of *Mytilus trossulus* and *Macoma balthica* to NE Europe it has been considered that Pacific taxa arriving through the Arctic route would have been already adapted to cold environments and thus were capable to colonize new habitats earlier than the resident Atlantic taxa from their southern refugia (Väinölä 2003, Johannesson & André 2006, Strelkov et al. 2007). A similar scenario could apply to the post-glacial invasion of the Pacific herring: these trans-Arctic invaders might have been better adapted to cold and to utilizing the crustacean plankton resources of northern marginal seas as soon as they were de-glaciated, compared to the native Atlantic herring (I).

The presence and duration of the pelagic larval stage (Burton 1983, Bohonak 1999, Hellberg et al. 2002, Riggins et al. 2011) and affiliation to a subtidal habitats (Buonaccorsi et al. 2002, Kelly and Palumbi 2010) have been found to correlate with genetic connectivity among populations in marine taxa. Regarding the inferred trans-Arctic dispersal, neither of these traits seems to be particularly linked to dispersal frequency. Only three of the invertebrates of which enough data were available showed deep inter-oceanic divergence without secondary dispersal: the barnacle *Semibalanus balanoides*, the horse mussel *Modiolus modiolus* and the sea star genus *Asterias*. Both *Modiolus* and *Asterias* should however be relatively good dispersers, with pelagic larval stages lasting up to several months (Barker & Nichols 1983, Coan et al. 2000, Roberts et al. 2011). Regarding subtidal habitat choice, the barnacle *S. balanoides* occupies shallower habitats than *Balanus crenatus*, which indeed has plausibly been engaged in more than one trans-Arctic invasion. However, also *M. modiolus* and *Asterias* are typically subtidal taxa. In addition, the vicariance pattern was detected in several deepwater benthic fishes (although by limited data) *Hippoglossoides robustus/H. platessoides* (this study), *Anarhichas* spp. (McCusker & Bentzen 2010) and *Gymnocanthus tricuspis/G. intermedius* (Yamazaki et al. 2013). One possible explanation is that these deepwater taxa have been relatively resistant to local extinctions and thus efficient in impeding new colonizations.

Adaptation to brackish water might also be related to trans-Arctic dispersal efficiency. Walters (1955) suggested that fish species that are adapted to a brackish water environment would be able to survive within Arctic refugia during the glacial periods. Such taxa might persist in big river estuaries, where the outflow kept the environment
more temperate. From such refugia they would have been able to follow the retreating ice sheets and be the first (and best adapted) arrivals in the deglaciated regions. Such adaptations plausibly played a crucial role during the Pleistocene glacial cycles. Hypothetically, the inferred inconvenient divergence time of *Zoarces* during the major glaciations (ca 1 Myr ago) could be explained by such Arctic refugia, considering the adaptation of *Zoarces viviparus* to brackish water. Similarly, such mechanisms might have been important for some of the euryhaline polychaetes, such as *Harmothoe imbricata* and *Glycera capitata*, both with evidence of several invasions. Nevertheless, the pattern is quite the opposite for the European flounder *Platichthys flesus*, which shows deep divergence and no signs of recent dispersal, despite that it can spend part of its life cycle in freshwater unlike other flatfish species (Kottelat & Freyhof 2007).

A successful colonization will probably always require a combination of many facilitating factors. Apparently, chance can also be important for the outcome. In echinoderms, for which the trans-Arctic dispersal seems to be an ongoing process, the success of long-distance invasions in the tropical oceans has been inferred to be of very stochastic nature (Palumbi 1996, McCartney et al. 2000, Addison & Hart 2005).
5. CONCLUSIONS AND FUTURE PROSPECTS

The amphi-boreal marine fauna has turned out to be taxonomically and genetically a much more complex system than hitherto believed. This is especially true in the seas close to the Arctic margins, where the glaciations have caused local extinctions and thus facilitated establishment of new invaders by vacating niches repeatedly. It is also significant in the sense that they are in general also rather species-poor areas. Molecular studies have revealed that taxa traditionally considered as single species in fact often are not coherent, regionally monophyletic entities, but may consist of elements representing two or more distinct genetic lineages, formed as a result of repeated trans-Arctic isolation-colonization cycles.

At the present, several examples exist of taxa or populations arisen from, or influenced by, the fusion of previously isolated lineages in secondary contacts of amphi-boreal sister taxa. These examples highlight the independent evolutionary potential of hybrid populations in boreal marine fauna and imply that also in this context, hybridizing genomes may in the future create new taxa, “synthetic species”, as has been demonstrated in corals (Frade et al. 2010), sculpins (Stemshorn et al. 2011), cichlids (e.g. Saltzburger et al. 2002) and possibly dolphins (Amaral et al. 2014). While the actual reverse speciation is not considered here, the amalgamation of genomes has confounded the systematic identities more regionally, in the marginal parts of the target basins. There seems to be some correlation between the levels of introgression and divergence time.

The hybridizing taxa in *Mytilus* and *Clupea* are judged to show independent histories of some 3.5 My, whereas for *Macoma* the vicariance was probably shorter (ca 2 Myr on paleontological and genetic grounds). For *Mytilus* and *Clupea*, the species identity is well retained in most part of the contact areas, while greater introgression has taken place only in local populations (Baltic Sea and Balsfjord; Strelkov & Väinölä 2011, II). In *Macoma balthica*, hybrid swarms in genetic equilibrium cover larger areas (Strelkov et al. 2007, Nikula et al. 2008).

Opportunities for dispersal through the Arctic are expected to increase in the future due to the warming climate (e.g. Greene et al. 2008, Vermeij & Roopnarine 2008, Miller & Ruiz 2014, Wisz et al. 2015). As a first sign of the possible coming trans-Arctic invasion, a Pacific diatom *Neodenticula seminae* was recorded in the Atlantic realm in the 1990’s for the first time in more than 800 000 years (Reid et al. 2007). Miettinnen et al. (2013) suggested the viable standing stock recorded during 2006–2008 would have originated from a pulse of Pacific surface water during the minimal sea ice cover in 2005 and 2007. Increased Arctic activity of subarctic cetaceans might also anticipate future invasions. These observations include trans-Arctic gene flow over the past millennium in the bowhead whale *Balaena mysticetus* (Alter et al. 2012), the Pacific ecotype killer whale *Orcinus orca* recorded in Newfoundland (mitogenome data; Morin et al. 2010) and the increased observations of humpback whales *Megaptera
Conclusions and Future Prospects

novaeangliae, fin whales Balaenoptera physalus, minke whales B. acutorostrata and killer whales Orcinus orca in the Chukchi Sea region (Clarke et al. 2013).

In the late 1980’s or the early 1990’s a fundamental reorganization of surface water circulation patterns occurred in the Arctic Ocean (Dickson et al. 2000), which caused a massive water fraction to enter the North Atlantic via the Canadian Archipelago, instead of the previous middle Atlantic route (Greene et al. 2008). Accordingly, at the present there seems to be a biogeographical range shift taking place in the NW Atlantic, where boreal plankton and fish are expanding southwards (Rose et al. 2000, Johns et al. 2001, Vilhjálmsson et al. 2005, Reid et al. 2007). On the contrary, in the NE Atlantic subtropical and temperate plankton and fish are shifting northwards (Beaugrand et al. 2002, Perry et al. 2005, Vilhjálmsson et al. 2005). These observations might suggest that the North American Atlantic coasts might be more affected by putative Pacific invaders than the European coasts.

The recent global warming resembles the early Holocene warming (Holocene Thermal Maximum), in the sense that both of the periods are linked to the Arctic Oscillation and the warming was initiated within the Pacific Ocean (Kaufman et al. 2004). However, during the Holocene Thermal Maximum the climate was balanced by the cooling effect of the remaining North American ice sheet, but during the future warming such balancing forces will not be present. Vermeij and Roopnarine (2008) argued that of the mere Bering Sea molluscs, at least 77 taxa have the potential to extend to the North Atlantic in the future decades without direct human assistance.

Trans-Arctic invasions should be a topical area of research in the near future. In order to attain an authentic view of the past invasion dynamics through the Arctic, wide-range sampling and international cooperation will be needed. This should also cover different taxonomical groups, since inclusion of genetic information across genera allows inferences on large-scale biotic events. An ideal approach would combine genetic and ecological data with geological, hydrographical and climatological information and utilize modern statistical and coalescence based methods designed for large datasets.
ACKNOWLEDGEMENTS

I would like to thank my supervisor Risto Väinölä for giving me the opportunity to work in this project and introducing me into the scientific way of thinking. You have taught me that solid inferences can only be made by truly knowing your data and that it is essential to understand what are the relevant questions to ask. You have always had time for me and never been in a rush when something has needed to be discussed.

During my PhD project, I have had the pleasure to work with several great biologists. Petr Strelkov, you’re such a genuine person. I’m really grateful for all of your input in the field, in the lab and beside the computer. Michael Hardman, I’m indebted for everything you have taught me about molecular lab work. Thank you also for the great fieldwork company and always keeping up the good spirit. I want to thank Dmitry Lajus; it was an honor for me to participate to the research on White Sea herring. Mikhail Daneliya, you have helped me in and outside the lab during this project, a big thank you! I want to thank Soili Stenroos and Alf Norkko for the guidance as members of my Thesis Advisory Committee. The reviewers of this thesis, Laura Kvist and John Wares, thank you for the valuable comments you have given to this work.

Special thanks to Dr. Michael L. Zettler and Alexander Darr for including me in your southern Baltic Sea expedition. I also want to thank Marina Katolikova, Kira Regel, Petteri Lehikoinen, Heini Kujala, Andrey Voronkov and Oddvar Skoggli for providing me irreplaceable help with sample collections.

All the colleagues at the zoological museum, thank you so much for all the help and friendship during these years. I especially want to thank Katja Nylund, I appreciate your input enormously. You are superb! During the years, there has been a varying but always wonderful group of PhD-students at the museum. The peer-support and joy this group has brought me has been essential, thank you! I want to thank the zoology unit director Aino Juslén and the director of museum Leif Schulman for being always interested in the well-being of the PhD students.

Sometimes I’ve doubted my choice of career, but luckily there’s been some great friends to convince me otherwise. Thank you Sanna, Maria, Sanni, Essi and others who I’ve come to know from Symbioosi, Heini, Karoliina, Laura L, Laura N, Meri, Minttu, Reetta, Sanna K, Sanna M, Sanna S, Suvi and others from Bio- (and anti-) sukeltajat and all you great friends from and around birding. Some of you have even taken the time to travel with me to collect samples for this thesis (Heini, we found at least three new species of *Hiatella*!), or searched bivalves on our holiday trips (Pepe, Heikki and Jonne, the time taken from birding yielded two new *Hiatella* species). Indeed, you all in your unique ways are pearls among people. Thanks for your example in science and in life. Thank you for taking me to places I would not have otherwise gone. Some friends have remained through the decades while life has taken us to different directions: Laura and
Marjut, you tolerated my interest to bugs and creatures throughout our childhood. Thanks for all the best times in the past and in the present.

Huge thanks to my family and to my extended Lehikoinen family. Marja ja Sakke, kiitokset kaikesta avusta vuosien varrella; kiitos näytteidenkeruuavusta, mutta ennen kaikkea siitä kaikesta muusta. Äiti, et ole koskaan pelännyt uusia asioita ja olet aina antanut minun ymmärtää että voin tehdä elämässä aivan mitä haluan. Kiitos että olet tukenut minua kaikissa valinnoissani. Olette rakkaita kaikki.

Pepe. The making of PhD was sometimes stressful, but I have only afterwards realized how much it also affected you. Still, as always, you supported and helped me in everything. It's hard to thank enough someone who is so essential part of one's life. You're the best, thank you for being you.

This thesis has been funded by the Academy of Finland (project grant #127471 to RV) and further personal grants from the Emil Aaltonen foundation, the Finnish Cultural Foundation, the University of Helsinki, the Oskar Öflund foundation, Societas pro Fauna et Flora Fennica and Kuopio Naturalists’ Society to HML. Funds for sampling trips and conferences were provided by the ASSEMBLE framework (grant agreement no. 227799), the Helsinki University and LUOVA (Doctoral Programme in Wildlife Biology Research, HY).
REFERENCES


Bouzat, J. L. (2000). The importance of control populations for the identification and


Trans-Arctic dispersals and the evolution of a circumpolar marine fish species complex, the capelin (Mallotus villosus). Molecular Ecology, 16, 5030-5043.


Huang, W., Takebayashi, N., Qi, Y. & Hickerson, M. J. (2011). MTML-MsBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. BMC Bioinformatics, 12, 1.


reconstruction and data-model comparison for the middle Pliocene. Global Ecology and Biogeography, 17, 432-447.


Summary


