Preface

The thesis is submitted in fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Faculty of Biosciences and Aquaculture (FBA), Nord University, Bodø, Norway. The original research presented in this thesis through five research papers was carried out on a period of three years, from 15.09.2014 to 25.10.2017. This research has been part of the following projects financed by the Norwegian Research Council: (1) "*Calanus* in the North Atlantic: species distribution and genetic population structure in space and time" (HAVKYST 216578, 234356 and part of FP7 EURO-BASIN - Grant Agreement: 264 933); (2) *COPPY*: "Fate of COPePod secondary production in a changing Arctic" (Polarprog. Project 227139, 2013-2016) and (3) "Isfjorden Marine Observatory System" (Polarprog. Project 246747, 2015-2018).

The project team consisted of the following members:

Marvin Choquet, MSc: PhD Student

Galice Hoarau, Professor, FBA, Nord University, Bodø, Norway: primary supervisor

Janne E. Søreide, Associate Professor, Department of Arctic Biology, The University Centre in Svalbard, Longyearbyen, Norway: co-supervisor

Marvin Choquet,

Bodø, October 2017

Committee

1. Opponent

Dr. Leocadio Blanco-Bercial

Bermuda Institute of Ocean Sciences

St. George’s, Bermuda

2. Opponent

Dr. Fredrika Norrbin

Department of Arctic and Marine Biology

Faculty of Biosciences, Fiskeries and Economics

UiT the Arctic University of Norway

Tromsø, Norway

Leader of the committee

Associate Professor Henning Reiss

Faculty of Biosciences and Aquaculture, Nord University

Bodø, Norway

I dedicate this thesis to the memory of my Grand-Mother,

who was the kindest person of my world

We must accept finite disappointment, but never lose infinite hope.

Martin Luther King, Jr.

Acknowledgements

"At some point, everything's gonna go south on you... everything's going to go south and you're going to say, this is it. This is how I end. Now you can either accept that, or you can get to work. That's all it is. You just begin. You do the math. You solve one problem... and you solve the next one... and then the next. And if you solve enough problems, you get to come home."

Mark Watney, in The Martian (Movie: 2015 – Dir. Ridley Scott)

This quote reminded me of the way I pursued my PhD, with the constant effort not to give up despite difficulties and failures. There are a lot of people who supported me over these 3 years, and who largely contributed, on different levels, to the final thesis. First, I acknowledge the Norwegian Research Council and Nord University for support of my research and position.

I would like to thank my main supervisor, Galice, for all the support he provided to me from the first day. Thank you for your positivity, your availability and all the daily scientific discussions, brainstorming and eternal flow of new research ideas, always full of enthusiasm. I also want to thank you for making me discover new activities when I first arrived in Norway, for example teaching me how to ski (despite the resulting "big headache" and blue eye, I still like it).

I would like to warmly acknowledge my co-supervisor, Janne, for her valuable advises and help in particular at the end of my PhD. Thank you for your kindness and your care about my workload. I am very grateful that you made me join the Polar Night cruise, which was my first cruise ever and I was so eager to be part of a cruise. Even though I was so seasick that from the second day already I thought I was going to die on board, and even though it was the polar night and I could not see nothing of the landscape, I still do not regret it.

For being the first persons taking great care of me when I arrived in Norway, thank you Irina and Alexander. Thank you Irina for sharing with me your experiences and wise advises.

Boris, we became best friends very fast, and I am grateful to all the things you made me discover, from hiking (and learning each mountain name) to fishing (and I definitely beat you with my 10kg's cod) and climbing. Thank you for making sure that I do not spend 100% of my time working only. I will remember of course the crazy parties in Sku Bare. Also, a big thought for you Marina, thank you for your kindness and caring so much for my well-being. I wish you both the best in Vancouver.

What a good idea you had, Chloé, to come and join the PhD-program at FBA in Bodø. After sharing the experience of a Master together in Portugal, I am glad we lived part of the PhD experience together again (with a lot of coffee/chocolate breaks). Thank you for your invaluable friendship, for the adventures we lived together, for your psychological support, for always being on my side.

Because you are the most entertaining person I know, thank you Prabhu! My fishes and plants also want to thank you a lot for keeping them alive and healthy. Once you introduced me as "the perfect man", I guess I would introduce you as "the coolest man". Thank you for being an awesome friend, for your constant willing to help, for being an amazing dancer, and for your positive attitude. Coming to work is less fun when you are not there.

Thank you Amalia for teaching me how to play the piano, I am so glad of all what I have learned with you. Also, thank you for drawing me as the copepod archer in your fantastic universe, I feel very proud about it.

Thank you Anusha and Martina for all your support in the lab, it was fun to share this working place with you. Thank you so much Morten for all your efforts in making sure I get all my samples on time. Thanks to all my fellows PhD and colleagues from the faculty for the nice time we had together, either sharing coffee breaks or hiking in the mountains. Thank you Jeanett for your strong support in the last days of my phD.

I would like to thank all the scientists with whom I have had the chance to collaborate, and who made me gain so much knowledge. I address a special acknowledgement to Arctos, for supporting me in my research and helping to organise the workshop.

Last but not least, I would like to thank my family for their love and support.

Table of contents

Preface **i**

Committee **ii**

Acknowledgements **iv**

Table of contents **vi**

List of figures **vii**

List of papers **viii**

Abstract **1**

1 Introduction **3**

1.1 Sentinels of climate change **3**

1.2 *Calanus* in the North Atlantic and Arctic **5**

Ecological and commercial importance 5

Calanus spp. life history 7

Distinct ecological niches 8

Species and body size matter 11

1.3 Species identification: challenges and advances **12**

1.4 Population genomics of zooplankton **14**

2 Objectives **17**

3 Main results **18**

4 General Discussion **21**

4.1 Misidentification of climate indicators, implications **21**

4.2 Fjords functioning and species ecological niches **22**

4.3 *Calanus* genus evolutionary history **23**

5 Conclusions and perspectives **25**

6 References **27**

List of figures

Figure 1
Calanus glacialis adult female 4

Figure 2
Life cycles of *Calanus* spp. 6

Figure 3
Calanus sp. stage CV, with a clearly visible lipid sack filling up its entire body cavity 7

Figure 4
Variability in life cycle duration depending on regions showing the flexibility of the *Calanus* species... 10

Figure 5
Pigmentation of *Calanus finmarchicus* and *C. glacialis* 13

Figure 6
Rearrangement of protein-coding genes order within mitochondrial genomes of *Calanus glacialis*, *C. sinicus* and *C. hyperboreus* 24

List of papers

Paper I **Choquet M**, Hatlebakk M, Dhanasiri AKS, Kosobokova K, Smolina I, Søreide JE, Svensen C, Melle W, Kwaśniewski S, Eiane K, Daase M, Tverberg V, Skreslet S, Bucklin A, Hoarau G.
Genetic redraws pelagic biogeography of *Calanus.*
Submitted to *Biology Letters*.

Paper II **Choquet M**, Kosobokova K, Kwaśniewski S, Hatlebakk M, Dhanasiri AKS, Melle W, Daase M, Svensen C, Søreide JE, Hoarau G.
Can morphology reliably distinguish between the copepods *Calanus finmarchicus* and *C. glacialis*, or is DNA the only way?
Submitted to *Limnology & Oceanography – Methods*.

Paper III **Choquet M**, Søreide JE, Burckard G, Skreslet S, Hoarau G.
No evidence of hybridization between the co-occurring copepods *Calanus finmarchicus* and *C. glacialis.*
Manuscript.

Paper IV **Choquet M**, Alves Monteiro HJ, Bengtsson-Palme J, Hoarau G. (2017).
The complete mitochondrial genome of the copepod *Calanus glacialis*.
Mitochondrial DNA Part B, 2:2, 506-507.
DOI: 10.1080/23802359.2017.1361357

Paper V **Choquet M** - Smolina I, Dhanasiri AKS, Kopp M, Jüterbock A, Sundaram AYM, Hoarau G.
Population genomics of zooplankton: how to deal with the challenges of large genomes in non-model species.
Manuscript.

Annex 1 Bucklin A, DiVito K, Smolina I, **Choquet M**, Questel J, Hoarau G, O'Neill R.

Population Genomics of Marine Zooplankton.

Chapter 19 in: Population Genomics: Marine Organisms

Submitted to Editors Om P. Rajora and Marjorie Oleksiak.

Abstract

Species of the genus *Calanus* dominate the zooplankton biomass in the North Atlantic and Arctic Ocean where they play a key role both as grazers and as prey for many commercially important species. *Calanus* species are frequently used as climate indicators due to their distinct environmental preferences. The overall goal of this thesis was to use currently existing molecular tools and to develop new ones in order to address critical ecological and evolutionary questions related to the genus *Calanus* in the North Atlantic and in the Arctic Ocean.

Species identification remains a challenge within the genus, especially between *C. finmarchicus* and *C. glacialis*. We used a set of nuclear molecular markers to accurately identify *Calanus* to species level to redefine their respective distributions. Molecular species identification revealed much wider and overlapping distributions than previously known for all four *Calanus* species inhabiting the North Atlantic and Arctic regions, questioning both the validity of previous morphological ID and the presumed ongoing range shifts forced by climate change. Furthermore, microsatellites data suggested that *C. glacialis* is more resident in the fjords compared to *C. finmarchicus*.

An assessment of the commonly used morphological ID criteria was conducted with the help of the molecular tools. None of the prosome length, the red pigmentation in antennules and in genital somite, nor the shape of the gnathobase and the structure of the fifth pair of swimming legs were 100% reliable for species identification. So far only DNA can reliably discriminate between species, although some of the morphological traits can be more useful in some regions than other. Misidentification is thus likely to be widespread.

In two Nordland fjords, *C. finmarchicus*, *C. glacialis*, *C. helgolandicus* and *C. hyperboreus* were co-occurring during winter-spring. Both *C. finmarchicus* and *C. glacialis* adult females and males were present simultaneously in the fjords, potentially allowing for hybridization. However, hybridization is not likely to occur as no hybrid

were found in the fjords and neither among >4400 individuals from the 83 distinct locations sampled in the North Atlantic and the Arctic Ocean.

In zooplankton, global lack of genomic resources available has hampered the development of population genomics approaches. Sequencing genomes, mitogenomes and transcriptomes is now critical. Thus, the mitogenome of *C. glacialis* was sequenced and annotated. Furthermore, using sequence capture enrichment, a set of ca. 100k SNPs was developed for *C. finmarchicus* and *C. glacialis.* These resources will be crucial to assess the connectivity between populations and species.

Overall, this thesis substantially advances our understanding of the *Calanus* species complex and dynamics in the North Atlantic and Arctic Oceans. The inclusions of molecular tools enabled us to fill important knowledge gaps regarding these key species. The new genomic resources developed will open the way to many new studies, to better understand the impact of climate change in *Calanus* at population, species and ecosystem level.

1 Introduction

1.1 Sentinels of climate change

Zooplankton species are considered to be good indicators of climate change because of their short life cycles, their sensitivity to temperature changes, their condition of being free floating organisms drifting with currents, and the fact that they are very scarcely exploited, which means that changes in their demography cannot be associated with trends in exploitation (Hays et al., 2005, Richardson, 2008). Indeed, zooplankton appears to be the fastest group of organisms to shift their distribution ranges in response to climate variability (Poloczanska et al., 2016).

However, zooplankton represents a huge variety of species from different orders and lineages, with ca. 6,000 species described within the holoplanktonic group alone (Wiebe et al., 2010). Therefore, some species might be more appropriate than others as climate indicators, depending on their specific role within the ecosystem and their sensitivity to the environment. In the North Atlantic and Arctic Oceans, species of the genus *Calanus* (Fig. 1) are extensively used as climate indicators.

Several elements make species of the copepod genus *Calanus* valid as climate change indicators. First, they are among the most studied zooplankton organisms, with ca. 100 scientific publications per year for the last 30 years (*Web of Science*). Second, they are very abundant and widely distributed in the North Atlantic and the Arctic Ocean, and they occupy key positions within their respective food webs. Third, their distribution is closely related to different water masses, and the species are considered as indicators of these respective water masses (Blachowiak-Samolyk, 2008, Bonnet and Frid, 2004, Conover, 1988, Daase and Eiane, 2007, Falk-Petersen et al., 2007, Jaschnov, 1970). Fourth, they have a relatively short life cycle, and are thus sensitive to unfavourable environmental conditions leading to lower reproduction success. For example, *Calanus* spp. time their spawning to ensure their offsprings get the best feeding and growth conditions. However, climatic changes affect the onset of the spring bloom (Reid et al., 2001), and this can result in a mismatch between the young

copepods and their food (Cushing, 1990). Such phenomenon is not only impacting the recruitment success of the *Calanus* species, but also cascades upwards in the food web (Beaugrand and Kirby, 2010, Søreide et al., 2010). Fifth, *Calanus* distributions are directly affected by climate change (Reid et al., 2003). Therefore, there has been considerable effort to document and model distributional changes of these species (Beaugrand et al., 2002, Chust et al., 2013, Kjellerup et al., 2012, Reygondeau and Beaugrand, 2011, Villarino et al., 2015).

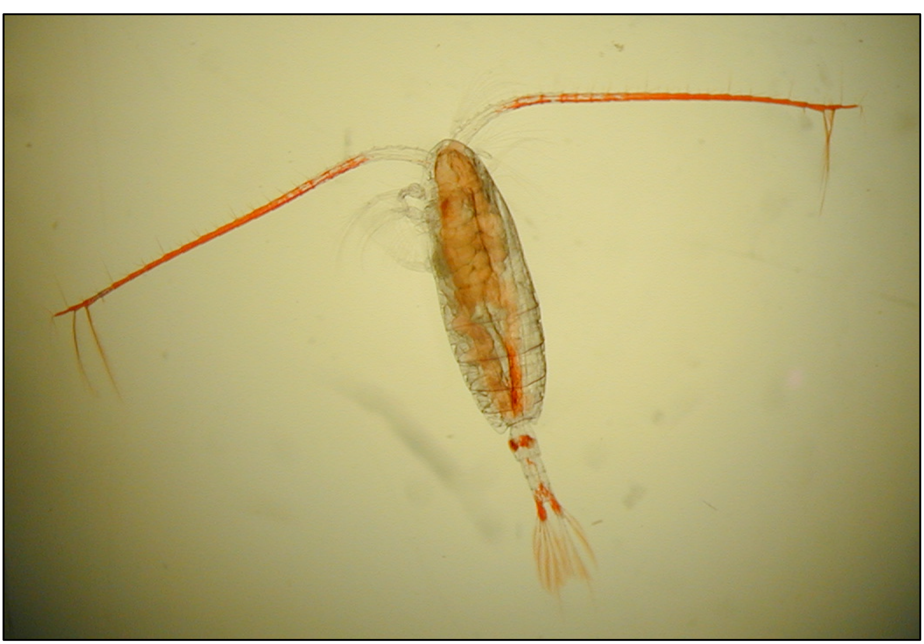


Figure 1: *Calanus glacialis* adult female. Photograph: S. Kwasniewski (Ny Ålesund, 2003).

1.2 *Calanus* in the North Atlantic and Arctic

Ecological and commercial importance

Species of the genus *Calanus* dominate the zooplankton biomass in the North Atlantic and Arctic Ocean (Blaxter et al., 1998, Conover, 1988, Fleminger and Hulsemann, 1977, Jaschnov, 1972, Kosobokova, 2012, Kosobokova et al., 2011, Søreide et al., 2008). They play a key role in marine pelagic food webs both as grazers and as prey for many commercially important species (Beaugrand and Kirby, 2010, Falk-Petersen et al., 2009, Jansen, 2016, Utne et al., 2012). *Calanus* spp. are able to store large amount of energy-rich lipids, converted from their phytoplankton diet (as reviewed in Lee et al., 2006), which makes them attractive food items for many organisms such as fishes (Gislason and Astthorsson, 2002), marine birds (Steen et al., 2007, Węsławski et al., 1999), marine mammals (Michaud and Taggart, 2007) and invertebrates (Falk-Petersen et al., 2002). Furthermore, by their ability to pack organic material into large fast-sinking faecal pellets, they are key drivers of the vertical export of material from the upper part of the water column to deeper layers (Riser et al., 2008, Wilson et al., 2008).

Recently, there has been a growing interest towards *Calanus* exploitation. A consensus has been obtained between Norwegian central authorities, funding bodies, R&D institutions and industry that zooplankton such as *Calanus* both can and indeed should be exploited. The company Calanus® AS (http://www.calanus.no/), based in Tromsø (Norway), has started to harvest this relatively large copepod in the Norwegian Sea following a quota limit suggested by the Institute of Marine Research (IMR, Norway). Currently 165,000 tons are withdrawn from the sea each year, used to make pills containing *Calanus* oil that represent a healthy source of omega 3 and helps to reduce symptoms associated to obesity and insulin resistance problem for people with diabetes type II. Harvested copepods are also used in feeds for aquaculture.

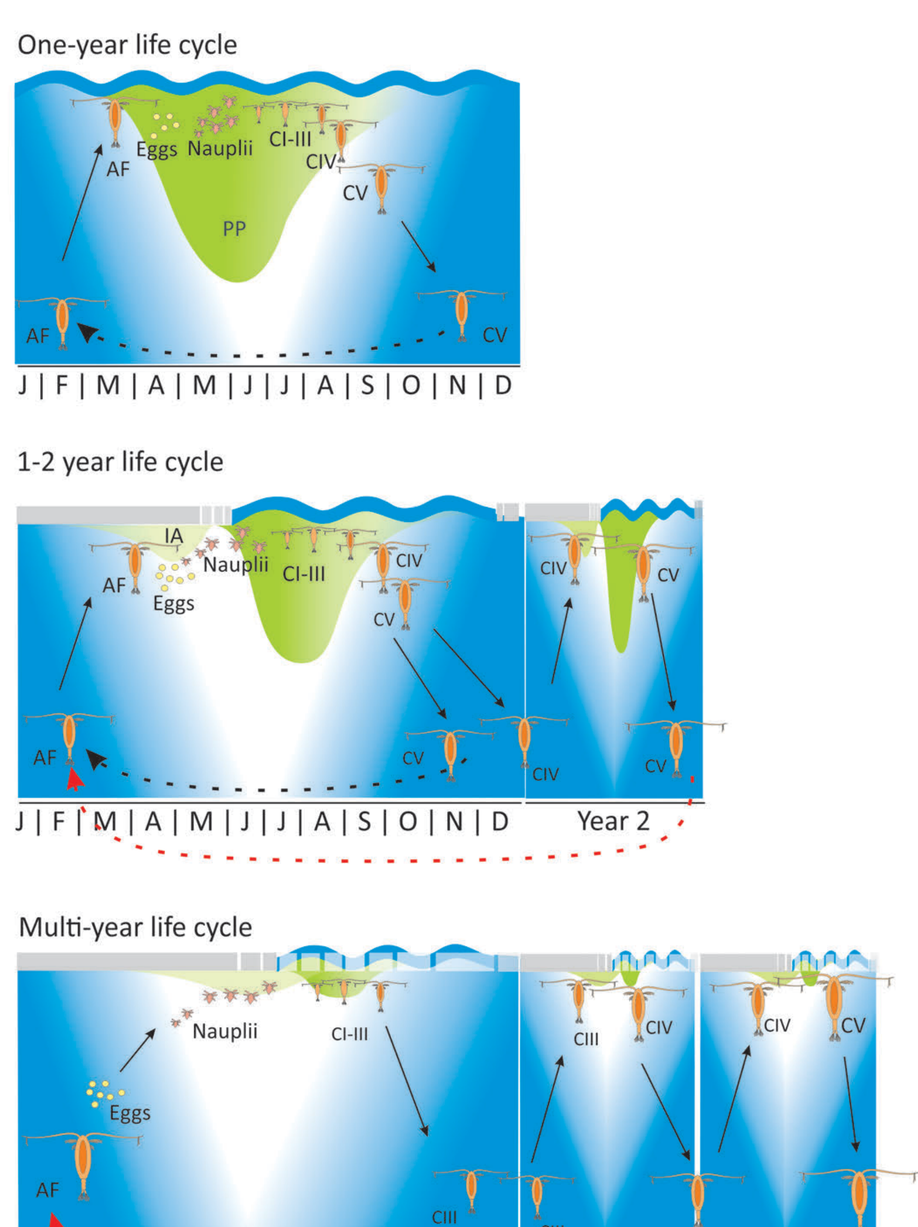


Figure 2: Life cycles of *Calanus* spp.. One-year life cycle (Top) is the most common for *C. finmarchicus* and *C. helgolandicus*. 1-2 years life cycle (Middle) is commonly displayed by *C. glacialis*. Multi-year life cycle (Bottom) is more common for *C. hyperboreus*. CI-CV: copepodite stage CI to CV; AF: adult female; IA: ice algae; PP: phytoplankton; Solid black arrows: seasonal migration; hatched red and black arrows connect the cycle. Illustration: Malin Daase.

Calanus spp. life history

Four *Calanus* species occur in the North Atlantic and the Arctic Ocean: *Calanus helgolandicus*, *C. finmarchicus*, *C. glacialis* and *C. hyperboreus*. They are morphologically very similar (Fleminger and Hulsemann, 1977, Frost, 1974,) and follow the same scheme of life development (Fig. 2). They are broadcast spawners and after hatching, the offspring follows 6 naupliar and 5 copepodite stages of development, before moulting to adulthood (=stage copepodite CVI). During their copepodite development, *Calanus* undergoes a diapause (Conover, 1988, Hirche, 1983, Madsen et al., 2001) usually in autumn-winter, when food is scarce. Depending on species and region, they migrate down to deeper sea layers as CIII, CIV or CVs, reducing their metabolic activity to a minimum (Hirche, 1997, Hirche, 1983) for 3 to 8 months.

The four species are predominantly herbivores (Falk-Petersen et al., 2009, Paffenhofer, 1976), but are able to switch to preys, such as protozooplankton (Levinsen et al., 2000) or even their own nauplii (Basedow and Tande, 2006, Bonnet et al., 2004) in case of phytoplankton shortage. *C. helgolandicus* has even been shown to feed on dead particles (Paffenhöfer and Strickland, 1970). The success of *Calanus* spp. can be explained by their ability to accumulate low-energy carbohydrates and proteins produced by phytoplankton and ice algae, which they convert into high-energy wax ester lipids (Falk-Petersen et al., 2009, Lee et al., 2006,). These lipids stores (Fig. 3) can allow their survival during periods of poor feeding or starvation (Gatten et al., 2013).

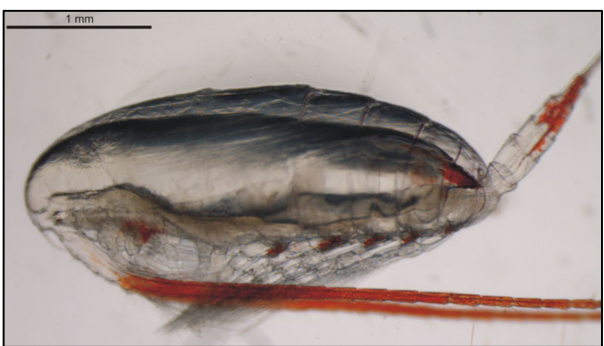


Figure 3: *Calanus* sp. stage CV, with a clearly visible lipid sack filling up its entire body cavity. Photograph: J. E. Søreide (Rijpfjorden, Aug. 2007, deep waters).

Distinct ecological niches

Although the four *Calanus* species have similar life histories, they differ in their life cycles (Fig. 4), on their preferred habitat, body size and lipid contents. *Calanus* are present everywhere in the North Atlantic and Arctic Ocean, but each species seems to have a distinct preferred habitat – a core distribution area with particularly high abundances.

Calanus helgolandicus distribution allegedly extends from the Mediterranean Sea to the North Sea (Barnard et al., 2004). The species is considered pseudo-oceanic (Helaouët and Beaugrand, 2007) because its distribution is centred over areas between 0 and 500 m depth, and thus very much influenced by bathymetry (Bonnet et al., 2005). Temperature also strongly drives *C. helgolandicus* distribution, associated with warm (9-20°C) temperate waters (Bonnet et al., 2005, Williams, 1985). *C. helgolandicus* may have up to 2 generations per year (Fig. 4), and is mostly dependent on the phytoplankton bloom as a resource to fuel its reproduction (income breeding strategy) (Planque and Fromentin, 1996). However, in cases of food restriction, reproduction can rely on lipid stores (capital breeding strategy), but egg production is then much less successful (Rey-Rassat et al., 2002).

Calanus finmarchicus is an oceanic species with its core distribution area in the Norwegian Sea and the Labrador Sea (Falk-Petersen et al., 2009), preferring colder temperate (estimated 4-12°C) waters (Jaschnov, 1961, Rees, 1957). *C. finmarchicus* is extensively used as an indicator of North Atlantic water masses (Helaouët and Beaugrand, 2007, Jaschnov, 1970, Jaschnov, 1966, Kwasniewski et al., 2003). Its ecological niche is well distinct from that of *C. helgolandicus* as *C. finmarchicus* seems better adapted to more variable and unpredictable environments (Helaouët and Beaugrand, 2007). However, recurrent records of its occurrence in the Arctic have so far been attributed to advection (Broms et al., 2009), and the hypothesis of local Arctic populations has been rejected as the environmental conditions are considered to be suboptimal for its recruitment (Diel and Tande, 1992, Hirche and Kosobokova, 2007, Melle and Skjoldal, 1998). The transport of individuals into the Barents Sea and the

Arctic Ocean represents a great loss for the North Atlantic system, but an important input of food for predators of the Arctic system (Hirche and Kosobokova, 2007). *C. finmarchicus* is mainly known as an income breeder (Richardson et al., 1999) that uses the phytoplankton bloom to fuel its maturation and spawning (Diel and Tande, 1992, Hirche, 1990, Plourde et al., 2001, Tande, 1982). However, in the Norwegian Sea and in the Barents Sea, spawning at lower rate before the bloom has also been reported (Hirche et al., 2001, Melle and Skjoldal, 1998). *C. finmarchicus* reaches adulthood within a year life cycle, and may have up to 2 generations per year (Broms and Melle, 2007, Matthews et al., 1978) in its southern distribution range (Fig. 4).

Calanus glacialis is regarded as an endemic Arctic species, with its main distribution located north of the polar front (Conover, 1988, Jaschnov, 1970, Jaschnov, 1961). The species is associated with shelves environments (Conover, 1988) and is present everywhere along the coasts in the Arctic (i.e. circumpolar), and in the White Sea. *C. glacialis* is sensitive to increases in sea temperatures and 5 to 6°C seems to be a threshold temperature for its well-being (Carstensen and Weydmann, 2012). At lower latitudes in the White Sea and Lurefjord (southern Norway) it migrates down to colder layers as soon as surface temperature starts increasing every year (Niehoff and Hirche, 2005, Pertsova and Kosobokova, 2010). *C. glacialis* is considered as an indicator of water masses of Arctic origins (Broms et al., 2009, Jaschnov, 1970, Kwasniewski et al., 2003, Unstad and Tande, 1991). *C. glacialis* has a mixed capital and income breeding strategy and 1 to 2 year life cycle. The mixed breeding strategy allows for more flexibility, helping *C. glacialis* to cope with the highly variable and unpredictable environmental conditions in the Arctic. When living in seasonal ice-covered shelf-seas, *C. glacialis* can utilize the ice algae bloom to fuel gonad maturation and spawning, allowing the offspring to benefit on the later occurring phytoplankton bloom (Daase et al., 2013, Søreide et al., 2010). In ice-free seas, where *C. glacialis* cannot benefit from a bimodal algal spring bloom with ice algae preceding phytoplankton, the species relies mainly on a capital breeding with fewer but more lipid-rich eggs to ensure recruitment (Hatlebakk, 2014).

Calanus hyperboreus is defined as a sub-Arctic and Arctic oceanic species (Broms et al., 2009, Conover, 1988) and is considered as an indicator of Arctic oceanic water masses (Conover, 1988, Jaschnov, 1970) with its distribution centre in the Greenland Sea (Hirche, 1991). Regular records in the northern Norwegian Sea have been documented and attributed to individuals advected through Arctic intermediate waters (Broms et al., 2009). Nonetheless, a local self-maintaining population of *C. hyperboreus* has also been described in Vestfjord (Melle and Skjoldal, 1998). *C. hyperboreus* is usually much bigger in size compared to all the other *Calanus* species, and this can be explained by its extended life cycle, which can last for 4-5 years (Falk-Petersen et al., 2009) (Fig. 4). Because of living in the central Arctic ocean, where phytoplanktonic production is lower and highly unpredictable, the species is almost exclusively a capital breeder (Conover and Siferd, 1993, Hirche and Niehoff, 1996).

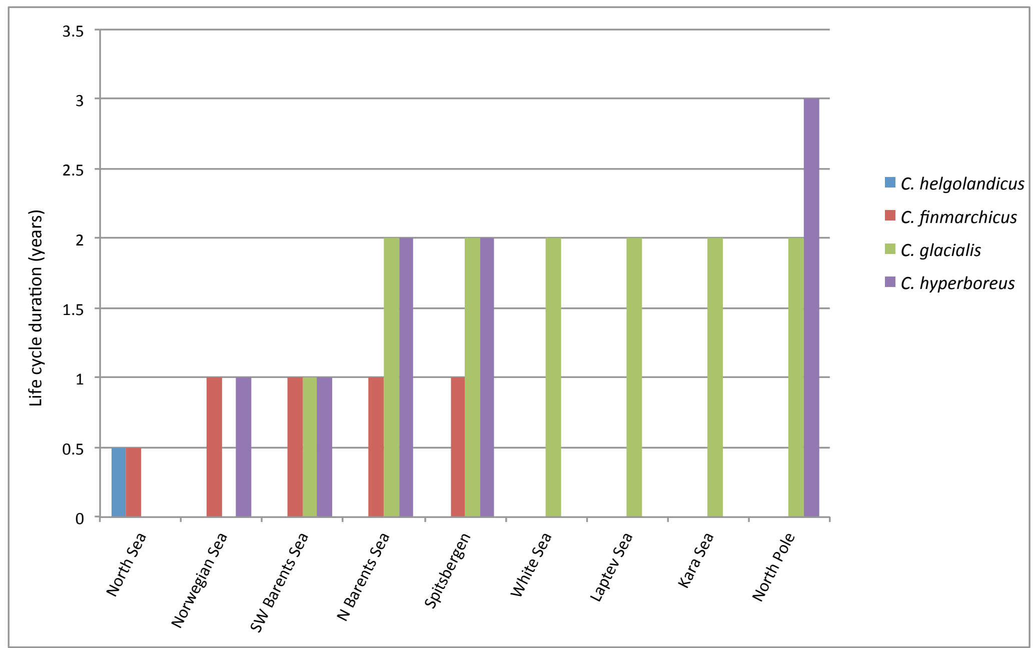


Figure 4: Variability in life cycle duration depending on regions showing the flexibility of the *Calanus* species. Illustration based on review of the followings: (Arnkværn et al., 2005, Bonnet et al., 2005, Broms and Melle, 2007, Falk-Petersen et al., 2009, Falk-Petersen et al., 1999, Hirche, 1997, Jaschnov, 1970, Jaschnov, 1961, Jaschnov, 1939, Kamshilov, 1955, Kosobokova, 1999, Kosobokova et al., 1997, Matthews et al., 1978).

Species and body size matter

These four *Calanus* species resemble each other both morphologically and in life history, but with some distinct differences when it comes to life cycle duration, reproduction strategy, size and environmental preferences. Longer life cycles result in adult individuals growing bigger in size, compared to the individuals that reach adulthood within a year. Thus, in zooplankton, size is a plastic trait that is mainly controlled by developmental time, which depends on temperature and food availability (Campbell et al., 2001, Huntley and Lopez, 1992). Indeed, developing slower, on a 2 years time scale (e.g. *C. glacialis*) or more (*C. hyperboreus*), enables the copepods to grow bigger and thus to accumulate more wax esters rich lipids (Falk-Petersen et al., 2009). *Calanus* lipid mass has been shown to be directly related to its body size (Vogedes et al., 2010). *Calanus* species complex has been described as a good example of interspecific Bergman cline because each of the species are significantly larger in the Arctic versus the more temperate waters (Leinaas et al., 2016). This confirms that the environment drives differences in size between species, with a bigger size being observed in unpredictable environments such as the Arctic Ocean. Body size of *Calanus* matters a lot for the rest of the ecosystem because of the food amount it represents. In Arctic, the little auk is performing a bi-modal foraging strategy, feeding partly on *C. glacialis*, closer to the shore and easier to reach, and partly on *C. hyperboreus*, further offshore (Steen et al., 2007). Despite the extra effort required to reach *C. hyperboreus*, the bird preys on it because it represents a much richer source of food per prey item (Steen et al., 2007).

Calanus body sizes are influenced by environmental conditions but not entirely determined by them. The extent to which *Calanus* can adjust its development time appears to be species specific (Fig. 4). For example, *C. finmarchicus* and *C. glacialis* overlap in body sizes in regions where they both complete their life cycle in one year (Gabrielsen et al., 2012); but in the Arctic Ocean where *C. glacialis* needs two years to complete its life cycle, resulting in larger individuals, we do not find any *C. finmarchicus* as large as *C. glacialis*. *C. finmarchicus* is almost exclusively an income

breeder and thus is not flexible enough to adjust its life cycle to two years, probably contributing to the absence of local recruitment in the Arctic (Hirche and Kosobokova, 2007, Kosobokova, 2012).

Furthermore, the concept of species within *Calanus* also matters for their ecosystems. For instance, *C. finmarchicus* and *C. helgolandicus* have similar developmental time duration and body size in the North Sea, but *C. finmarchicus* peaks in abundance in spring, while *C. helgolandicus* reaches its abundance maximum later in the summer. Juveniles of cods used to feed on abundant *C. finmarchicus* in spring to ensure successful recruitment. In the mid-1980's, when sea surface temperature triggered the progressive substitution of *C. finmarchicus* by *C. helgolandicus* in the North Sea, it resulted in the collapse of the cod population because of a mismatch between cod juveniles and their preys (Beaugrand et al., 2003).

1.3 Species identification: challenges and advances

Identification of species within the zooplankton has always been challenging because most of the taxa are very small and often lack well-expressed diagnostic features, particularly for the congeneric species (Aarbakke et al., 2011, Bailey et al., 2015, Frost, 1989, Frost, 1974, Jaschnov, 1957, Rees, 1949). As potential diagnostic characteristics may take hours to examine, even for experienced taxonomists, variable traits such as body size are thus often used for routine based identification. Larval stages and young developmental stages of zooplankton organisms are usually not identified at the species level (McManus and Katz, 2009) because of their small size.

Species of the genus *Calanus* are morphologically very similar (Fig. 5) and their identification has always been a challenge (Conover, 1988, Fleminger and Hulsemann, 1977, Frost, 1974). Despite the extensive literature, a limited number of diagnostic characters have been described. The most common method relies on the prosome length, but as size varies geographically with temperature and food condition (Fig. 4), it often leads to misidentification (Gabrielsen et al., 2012, Lindeque et al., 2006, Parent

et al., 2011). More complex characters have been described as species-specific for *Calanus* (Beklemishev, 1959, Jaschnov, 1955) but are rarely used because of the complexity of their examination, which requires both time and taxonomic skills. Recently, the pigmentation of the antennules and genital somite has been proposed as diagnostic character to differentiate between *C. finmarchicus* and *C. glacialis* (Nielsen et al., 2014) (Fig. 5). However, this character has only been tested for adult females from a single geographic location, and requires individuals to be alive.

Figure 5: Pigmentation of *Calanus finmarchicus* and *C. glacialis*.
Calanus finmarchicus with pale antennules and genital somite (top) and *C. glacialis* with red antennules and genital somite (bottom). Photograph: S. Kwasniewski (Ny Ålesund, 2003).

Therefore, in an effort to accurately distinguish these species, different molecular tools have been developed (Lindeque et al., 1999, Parent et al., 2012, Provan et al., 2009, Smolina et al., 2014). However, they still remain insufficiently used for species identification. In the context of climate change, *Calanus* species are expected to shift the distribution of their populations (Chust et al., 2013, Kjellerup et al., 2012, Wassmann et al., 2015, Wassmann et al., 2011), as it has been documented in the

1980's in the North Sea (Beaugrand, 2004) with *C. finmarchicus* being replaced by *C. helgolandicus*. It is expected that, via a process of "Atlantification" of the Arctic regions (Falk-Petersen et al., 2007, Wassmann et al., 2006), conditions in Arctic will become more and more favourable for *C. finmarchicus* to become the prevalent species there, replacing *C. glacialis* (Reygondeau and Beaugrand, 2011, Slagstad et al., 2011). Thus, areas where species co-occur are likely to increase in the future. In addition, it has been proposed that species within the *Calanus* genus may hybridize (*C. finmarchicus* and *C. glacialis*) (Parent et al., 2012). Therefore, in this context, correct species identification of *Calanus* spp. is critical if we want to use their distribution changes to detect impacts of climate change.

The development of molecular-based species identification methods to ensure reliable species recognition is becoming more and more common, mostly due to the availability of new powerful sequencing technologies with decreasing costs. For example, barcoding approaches rely on the comparison of a single locus (e.g. 16S, 18S, COI, 28S), present in every species but with species-specific variations (e.g. Bucklin et al., 2011, Bucklin et al., 2010a, Bucklin et al., 2010b, Bucklin et al., 2007, Grant and Linse, 2009, Strugnell and Lindgren, 2007, Trivedi et al., 2016).

The most important limitation of mtDNA barcodes is that they cannot be used to track hybridization (because of uniparental inheritance). Furthermore, up-scaling the number of individuals to analyse can be costly and time consuming. Thus, new nuclear markers have been developed for *Calanus* species identification based on partial transcriptome and genomes (Smolina et al., 2014). These markers offer an easier, faster and cheaper alternative to common barcodes.

1.4 Population genomics of zooplankton

Despite being extremely well studied, some important knowledge gaps remain for *Calanus* and other ecologically important zooplanktonic taxa. Indeed, little is known

about the dynamics of populations of zooplankton species, in terms of gene flow and connectivity. Such knowledge is, however, essential to understand the distribution of the species, and to predict responses to climate change. In the case of *Calanus* spp., several studies have used molecular markers to try to define the pattern of genetic structure between populations (Bucklin, 2000, Bucklin et al., 2000, Bucklin et al., 1996, Kann and Wishner, 1996, Nelson et al., 2009, Provan et al., 2009, Unal and Bucklin, 2010, Weydmann et al., 2016, Yebra et al., 2011). However, these studies were based on a limited number of molecular markers and individuals, and therefore contradictory results were obtained depending on studies and type of markers used. The question of whether *Calanus* species populations are genetically differentiated is thus still open.

With the recent rapid development of next-generation technologies of sequencing, it is now possible to assess the genetic structure between populations using tens of thousands of SNPs (Single Nucleotide Polymorphisms) distributed all over the genome (Helyar et al., 2011, Reitzel et al., 2013). This is a complete change of scale compared to previous studies using a limited number of molecular markers to compare populations. Having enough markers that represent the diversity of a whole genome will allow detecting more subtle structure. Furthermore, on-going developments in the genomics field bring new tools to address many important ecological questions, and increase our ability to predict community responses to climate change (Hofmann et al., 2005).

However, next-generation sequencing technologies have remained scarcely used in the zooplankton, mainly due to the global lack of prior genetic/genomic knowledge for most of these organisms. In zooplankton, only very few species can be considered as model species, with their genome fully sequenced and annotated (e.g. Denoeud et al., 2010, Madoui et al., 2017, Moroz et al., 2014, Ryan et al., 2013). One of the main challenges for many zooplankton species is their large genome sizes. For example, *Calanus* genomes are particularly large in size, estimated to 6.5 Gb for the haploid genome of *C. finmarchicus*, 10.5 Gb for *C. helgolandicus*, and 12.5 Gb for *C. glacialis* and *C. hyperboreus* (McLaren et al., 1988). In comparison, the human genome is 2 Gb

large. This large size may be explained by the presence of duplications within the genome. For many studies such as investigating the population structure of a particular non-model species with a large genome, sequencing of the complete genome is often unnecessary and would increase the cost and the complexity of the study (Narum et al., 2013). Different methods of genotyping-by-sequencing offer an alternative allowing us to characterize thousands of molecular markers selected via reduced-representation protocols (reviewed in Crawford and Oleksiak, 2016).

(For a review on the use of genomics in marine zooplankton organisms, see Annex 1)

2 Objectives

The overall goal of this thesis was to use currently existing molecular tools and to develop new ones in order to address critical ecological and evolutionary questions related to the key zooplankton genus *Calanus* in the North Atlantic and the Arctic Ocean. This was accomplished through the following objectives:

1) Redefine the distributions of the four *Calanus* species in the North Atlantic and Arctic Oceans using molecular tools (**Paper I**)

2) Evaluate the potential of morphological characters to distinguish between *Calanus* species, using a molecular-based approach (**Paper II**)

3) Study the potential of hybridization among the co-occurring *Calanus* species (**Papers I, III**)

4) Develop new genomic resources to address key ecological and evolutionary questions in *Calanus* (**Papers IV, V**)

3 Main results

Genetics redraws pelagic biogeography of Calanus

Objective 1: Contributions from Paper I

We used molecular markers to accurately identify *Calanus* to species level from samples gathered from most of the North Atlantic and the Arctic oceans, to redraw the actual distributions of the four species *C. helgolandicus*, *C. finmarchicus*, *C. glacialis* and *C. hyperboreus* in this area. The species distributions found were more extended and overlapping between species compared to what was described before. *C. helgolandicus*, the most temperate of the four species, was found as far North as 70° North and large proportions of the Arctic *C. hyperboreus* were found as far south as 60°N (Oslo, Norway). The most striking finding, however, was the high proportions of *C. glacialis* together with *C. finmarchicus*, in several fjords all along the Norwegian coast (from 60°N). We performed a population structure study, based on 10 microsatellites markers, and detected genetic differentiation within *C. glacialis* fjords populations, but not within *C. finmarchicus* populations. These results suggest that *C. glacialis* is more a resident species of the fjords than *C. finmarchicus*. To understand why *C. glacialis* fjord populations have not been detected before, we tested the accuracy of the most common morphological method used to distinguish between *C. finmarchicus* and *C. glacialis*, the prosome length. In Skjerstadfjord, our results showed that prosome length overlapped completely between the two species, preventing *C. glacialis* from being correctly identified.

Can morphology reliably distinguish between the copepods Calanus finmarchicus *and* C. glacialis*, or is DNA the only way?*

Objective 2: Contributions from Paper II

We used a set of molecular markers developed for *Calanus* species identification to evaluate the validity of different morphological characters described in literature as species diagnostic between *C. finmarchicus* and *C. glacialis*. Prosome length, redness of antennules and genital field, structures of gnathobase and fifth pair of swimming legs were tested on sets of molecularly characterised individuals. We found a strong variability of the different characters depending on regions, and noticed some areas where they seemed more species diagnostic than in others (northernmost areas, and areas of allopatry). This indicates that none of the morphological characters can be used to identify species with 100% reliability. We provided some recommendations about how carefully the morphological characters should be used, if used at all, depending on regions. Numerous past studies may have overlooked one or the other species, because of the use of morphological ID alone, and our current knowledge on this genus may have been strongly impacted.

Potential for hybridization among *Calanus* species living in sympatry

Objective 3: Contributions from Papers I and III

We used a set of nuclear molecular markers (InDels) developed to track the putative hybridization between *C. finmarchicus* and *C. glacialis.* We followed the stage and species composition of *Calanus* spp. in two boreal fjords along the Norwegian coast during the reproductive period (winter-spring). Molecular identification confirmed the co-occurrence of four *Calanus* species in these two fjords: *C. helgolandicus*, *C. finmarchicus*, *C. glacialis* and *C. hyperboreus,* with *C. helgolandicus* being the less abundant. In January, the co-occurrence of males and females of *C. glacialis*, *C. finmarchicus* and *C. hyperboreus* suggested that all three species reproduce on site.

Also, males and females of both *C. finmarchicus* and *C. glacialis* occurred simultaneously in both fjords, potentially leading to hybridization. However, we did not detect any hybrid amongst the 1,497 individuals genotyped (Paper III). Furthermore, using the same molecular markers, we also did not find any hybrid amongst 4,434 individuals from 83 locations across the North Atlantic and the Arctic Oceans (Paper I).

Development of new genomic resources to address key ecological and evolutionary questions in *Calanus*

Objective 4: Contributions from Papers IV and V

We followed three different approaches in order to develop genomic resources for future evolutionary/population genomics investigations. First, we sequenced and assembled the whole mitochondrial genome of *C. glacialis* using high-throughput sequencing. We annotated the sequence and used it to reconstruct a phylogenetic tree including the closest copepods species for which similar coding-gene regions were available. This sequence represents the longest mitochondrial genome reported in marine zooplankton so far, with 20,674 bp (Paper IV).

We then experimented with two methods of genome reduction, RAD-seq and targeted resequencing (or sequence capture enrichment). Given the genome size, the RAD-seq approach was only moderately successful (1,871 high quality SNPs with sufficient coverage identified in *C. finmarchicus*). However, the targeted resequencing approach allowed us to characterize 140,859 SNPs in *C. finmarchicus* and 115,928 SNPs in *C. glacialis*, of high quality and sufficient coverage. This set of SNPs represents a powerful tool to assess the genetic differentiation within populations but also to investigate, at the genome level, the putative porosity of species boundaries.

4 General Discussion

4.1 Misidentification of climate indicators, implications

The new distribution ranges of *Calanus* spp. uncovered by molecular analysis in the North Atlantic and Arctic are divergent from the previously described distributions of species, based on morphological identification (Paper I). The divergence between the morphological-based versus the molecular-based species distributions is at least partly due to species misidentification (Papers I and II; Gabrielsen et al., 2012). Indeed, we showed that none of the morphological characters used to discriminate between *C. finmarchicus* and *C. glacialis* were 100% reliable (Paper II). It seems that in regions where *C. finmarchicus* and *C. glacialis* coexist during their entire life cycle, it is harder or even impossible to differentiate them using morphology. With the on-going climate change, it is expected that *Calanus* species distributions will change and overlap even more (Slagstad et al., 2011), and then we may not be able to detect these changes if we keep using only morphology. An accurate understanding of *Calanus* species current distribution is crucial to track the effects of climate change on ecosystems (Richardson, 2008).

Importantly, we cannot exclude that the extended ranges revealed by DNA could result from on-going shifts that have remained cryptic due to the inaccuracy of morphological ID tools. Considering the importance of *Calanus* range shifts for our understanding of climate change impacts on pelagic ecosystems, our results question the current predictions and models. It is therefore critical to tease apart the respective effects of morphological misidentification from on-going range shifts and this will require a thorough reassessment of historical distributions using molecular tools.

4.2 Fjords functioning and species ecological niches

We reported high proportions of *C. glacialis* and *C. hyperboreus* in several Norwegian fjords, co-occurring with *C. finmarchicus* (Paper I). In Skjerstadfjord and Mistfjord, four species co-occur, with *C. finmarchicus*, *C. glacialis* and *C. hyperboreus* successively dominating the zooplankton in terms of abundance in Skjerstadfjord, while *C. glacialis* was the dominating species in Mistfjord during the whole period studied (Paper III). The three most abundant *Calanus* species showed to have adult females and males present at about the same time in both fjords (Paper III), which potentially allow for hybridization. However, the temporal resolution of our sampling may not have been fine enough to detect subtle differences in species-specific timing of females and males. To study the gonad development and in more detail the actual mating will significantly improve our biological understanding of why hybridization is not likely to occur between sibling species living in sympatry. Until now, *C. finmarchicus* has been considered as the main dominating species of the zooplankton assemblage in Norwegian fjords, and it is assumed to be advected in and out of fjords seasonally (Skreslet and Rød, 1986). The study on population structure conducted in Paper I suggests that this mechanism may not apply to *C. glacialis*, which questions our understanding of fjords circulatory systems and advective forces. One possible explanation is that *C. glacialis* is able to avoid the early summer flushing of fjords by going in deeper cold-water layers, as described in the White Sea and the Lurefjord (Niehoff and Hirche, 2005, Pertsova and Kosobokova, 2010) as a mechanism for the species to survive warm surface temperatures. Such hypothesis would suggest that an important component of the zooplankton community in fjords might have been overlooked in previous studies where only *C. finmarchicus* was considered.

During our winter-spring investigation of Skjerstadfjord and Mistfjord in Nordland, we found females and males of *C. finmarchicus, C. glacialis* and *C. hyperboreus* co-occurring. Despite such extensive sympatry we did not detect any hybrid (Papers I and III). Hybridization cannot be totally ruled out though, but if it happens it is likely to be very rare (or in the past).

Even though the different *Calanus* species are all successfully sharing similar habitats, such as the Norwegian fjords, it seems that each species occupies distinct ecological niches, but these specific niches with their limits remain to be investigated. We should therefore be careful about the growing development of fish farms at the entrance of many fjords, and the increase in the harvesting of *Calanus* for commercial purposes. These activities have started originally with the assumption that *C. finmarchicus* is the only species living in the fjords and that its populations are renewed each year by an offshore stock of individuals. Now, knowing that isolated populations of *C. glacialis* could also inhabit these fjords, I highly recommend re-assessing the state of knowledge on the trophic relationships and roles of each species, using molecular tools for species identification before further exploitation.

4.3 *Calanus* genus evolutionary history

The development of new genomic resources for *Calanus* such as a mitochondrial genomes and SNPs are crucial to address key ecological and evolutionary questions. The sequenced and annotated mitochondrial genome of *C. glacialis* (Paper IV) constitutes the third mitogenome publicly available for *Calanus* genus, along with *C. hyperboreus* (Kim et al., 2013) and *C. sinicus* (Minxiao et al., 2011). Within the same genus, these mitogenomes are quite different from each other, with large rearrangements of the genes order (Fig. 6). In marine zooplankton, such variability is common (Ki et al., 2010, Kohn et al., 2012, Marlétaz et al., 2017, Pett et al., 2011), compared to terrestrial organisms, and may reflect a strong evolutionary potential of marine zooplankton species.

However, the phylogenetic tree (16S) in paper I (Supp. 2) shows very low intraspecific variation in *C. glacialis*. Indeed, despite the sequencing of 138 individuals covering the North Atlantic and Arctic regions, we only found 4 haplotypes in *C. glacialis.* The other *Calanus* species were much more variable (*C. finmarchicus*: 10

haplotypes / 91 individuals; *C. helgolandicus*: 8 haplotypes / 26 individuals; *C. hyperboreus*: 14 haplotypes / 33 individuals). This low variability for *C. glacialis* is surprising given the extended geographical range of the species, and the fact that we detected some distinct differentiation between fjords populations (microsatellites) (Paper I). As shown in Fig. 6, there is a strong inter-specific genetic variability within *Calanus*, but a particularly low intra-specific variability for *C. glacialis*, questioning its evolutionary history compared to its congeneric species. Interestingly, the other Arctic species, *C. hyperboreus* appears to have the highest intra-specific diversity.

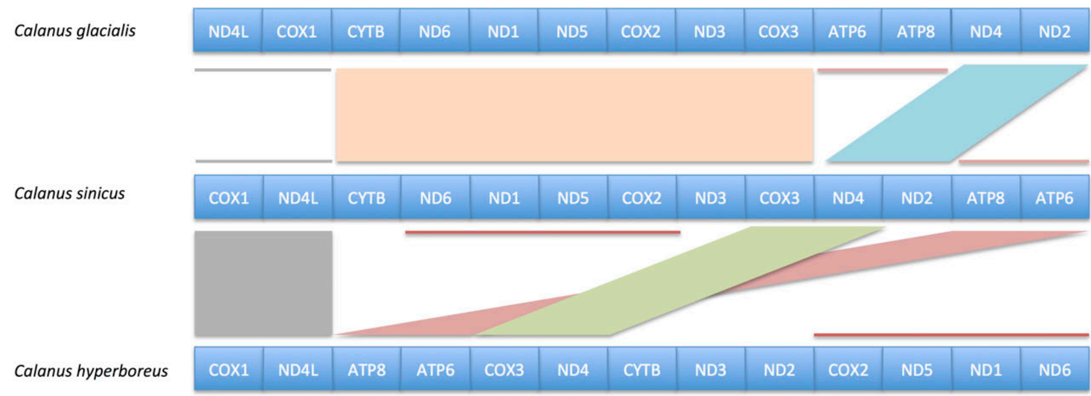


Figure 6: Rearrangement of protein-coding genes order within mitochondrial genomes of *Calanus glacialis*, *C. sinicus* and *C. hyperboreus*. Rectilinear shapes show genes for which the order is conserved between the species; lines indicate genes with the same sequence but in reverse order between the species.

The set of SNPs developed in Paper V will also contribute to further understanding of *Calanus* evolutionary history. The method can be extended to all four species and provide a genome wide picture of the evolutionary relationships within the genus, including putative introgression.

5 Conclusions and Perspectives

The present thesis illustrated the necessity to integrate molecular tools in classical biological oceanography. In this thesis, I combined ecological and molecular approaches to establish a new solid knowledge base for *Calanus,* which will open up for a wide variety of innovative new important research on *Calanus*. The *Calanus* complex has been widely studied, but knowledge gaps are still numerous. Species distributions have been well documented, but molecular markers have shown the weakness and bias of solely relying on morphology. Now that the species distributions are more accurately described, many questions remain. What are the drivers of these distributions? What makes the genus so successful in the North Atlantic? Genomics will be key to answering such questions and bring new elements to understand the *Calanus* complex in its whole. In this context, the sequencing of a genome of *Calanus* would be an important starting point despite the challenges of working with such large genomes.

The molecular method that we have developed (Papers I and II) is fast, easy to use and cheap. It does not require previous knowledge of genetics, or costly equipment. Moreover, our method only uses the antennules of the specimens. It is thus essential that such molecular method should be part of the routine species identification. Furthermore, in order to evaluate the response of *Calanus* spp. to climate change, it is critical to start looking at historical collections. The main challenge is that most historical samples have been preserved almost exclusively in formalin, and almost not in ethanol. However, the molecular identification of *Calanus* species from formalin-preserved samples has been tested and is possible (I. Smolina, personal communication) but remains to be carried-on on a large scale.

In Norwegian fjords, given the co-occurrence of the four *Calanus* species, we now need to better understand their place and role in the fjords ecosystem. Following their phenology throughout the year would be a good starting point to identify the different species life histories and niches.

The contrasted evolutionary histories among the *Calanus* species suggested by the 16S diversity should be investigated further. Newly developed mitogenome of *C. glacialis* (Paper IV) and available mitogenome of *C. hyperboreus* should be used to resolve and compare the phylogeography of the two species. Additional sequencing of *C. helgolandicus* and *C. finmarchicus* mitogenomes will allow us to investigate the history of the *Calanus* genus in the North Atlantic and Arctic, which will significantly improve our understanding of *Calanus* species specific distribution patterns.

The set of SNPs obtained from the DNA capture-enrichment based method (Paper V) may be used to look for introgression and porosity of species boundaries, examine and finally answer the questions on the population genetic structure of *C. finmarchicus* and *C. glacialis*.

More generally, we really need to reinforce the cooperation between ecology and genomics, in order to address the questions that cannot be answered by one discipline alone. Regarding species identification, we have to make this “successful marriage” of morphological and molecular methods as stressed in (McManus and Katz, 2009). Further, time is overdue to systematically collect zooplankton samples in ethanol and to generate “genomic friendly” time series. In a few years, such collection will be invaluable to evaluate the effects of anthropogenic changes on marine ecosystems.

6 References

Aarbakke, O.N.S., Bucklin, A., Halsband, C. & Norrbin, F. (2011). Discovery of *Pseudocalanus moultoni* (Frost, 1989) in Northeast Atlantic waters based on mitochondrial COI sequence variation. *Journal of Plankton Research,* 33**:** 1487-1495.

Arnkværn, G., Daase, M. & Eiane, K. (2005). Dynamics of coexisting *Calanus finmarchicus*, *Calanus glacialis* and *Calanus hyperboreus* populations in a high-Arctic fjord. *Polar Biology,* 28**:** 528-538. DOI 10.1007/s00300-005-0715-8.

Bailey, J., Durbin, E.G. & Rynearson, T. (2015). Species composition and abundance of copepods in the morphologically cryptic genus *Pseudocalanus* in the Bering Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*.

Barnard, R., Batten, S., Beaugrand, G., Buckland, C., Conway, D., Edwards, M., Finlayson, J., Gregory, L., Halliday, N. & John, A. (2004). Continuous plankton records: Plankton atlas of the North Atlantic Ocean (1958-1999). II. Biogeographical charts. *Marine Ecology-Progress Series***:** 11-75.

Basedow, S.L. & Tande, K.S. (2006). Cannibalism by female *Calanus finmarchicus* on naupliar stages. *Marine Ecology Progress Series,* 327**:** 247-255.

Beaugrand, G. & Kirby, R.R. (2010). Climate, plankton and cod. *Global Change Biology,* 16**:** 1268-1280. DOI 10.1111/j.1365-2486.2009.02063.x.

Beaugrand, G. (2004). The North Sea regime shift: Evidence, causes, mechanisms and consequences. *Progress in Oceanography,* 60**:** 245-262. DOI http://dx.doi.org/10.1016/j.pocean.2004.02.018

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S. & Reid, P.C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature,* 426**:** 661-664.

Beaugrand, G., Reid, P.C., Ibanez, F., Lindley, J.A. & Edwards, M. (2002). Reorganization of North Atlantic marine copepod biodiversity and climate. *Science,* 296**:** 1692-1694.

Beklemishev, K. (1959). On the anatomy of masticatory organs of Copepoda. Report 2: The masticatory edge in mandibles of certain species of Calanidae and Eucalanidae. *Trudy Instituta Okeanologii,* 30**:** 148-155.

Blachowiak-Samolyk, K. (2008). Contrasting zooplankton communities (Arctic vs. Atlantic) in the European Arctic marginal ice zone. *Oceanologia,* 50**:** 363-389.

Blaxter, J.H., Douglas, B., Tyler, P.A. & Mauchline, J. (1998). *The biology of calanoid copepods*: Academic Press.

Bonnet, D., Richardson, A., Harris, R., Hirst, A., Beaugrand, G., Edwards, M., Ceballos, S., Diekman, R., López-Urrutia, A. & Valdes, L. (2005). An overview of *Calanus helgolandicus* ecology in European waters. *Progress in Oceanography,* 65**:** 1-53.

Bonnet, D. & Frid, C. (2004). Seven copepod species considered as indicators of water-mass influence and changes: results from a Northumberland coastal station. *ICES Journal of Marine Science,* 61**:** 485-491.

Bonnet, D., Titelman, J. & Harris, R. (2004). *Calanus* the cannibal. *Journal of Plankton Research,* 26**:** 937-948.

Broms, C., Melle, W. & Kaartvedt, S. (2009). Oceanic distribution and life cycle of *Calanus* species in the Norwegian Sea and adjacent waters. *Deep Sea Research Part II: Topical Studies in Oceanography,* 56**:** 1910-1921. DOI 10.1016/j.dsr2.2008.11.005.

Broms, C. & Melle, W. (2007). Seasonal development of *Calanus finmarchicus* in relation to phytoplankton bloom dynamics in the Norwegian Sea. *Deep Sea Research Part II: Topical Studies in Oceanography,* 54**:** 2760-2775.

Bucklin, A., Steinke, D. & Blanco-Bercial, L. (2011). DNA barcoding of marine metazoa. *Annual Review of Marine Science,* 3**:** 471-508.

Bucklin, A., Hopcroft, R.R., Kosobokova, K.N., Nigro, L.M., Ortman, B.D., Jennings, R.M. & Sweetman, C.J. (2010a). DNA barcoding of Arctic Ocean holozooplankton for species identification and recognition. *Deep Sea Research Part II: Topical Studies in Oceanography,* 57**:** 40-48. DOI 10.1016/j.dsr2.2009.08.005.

Bucklin, A., Nishida, S., Schnack-Schiel, S., Wiebe, P.H., Lindsay, D., Machida, R.J. & Copley, N.J. (2010b). A census of zooplankton of the global ocean. *Life in the World's Oceans: Diversity, Distribution, and Abundance, edited by: McIntyre, A***:** 247-265.

Bucklin, A., Wiebe, P.H., Smolenack, S.B., Copley, N.J., Beaudet, J.G., Bonner, K.G., Farber-Lorda, J. & Pierson, J.J. (2007). DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). *Journal of Plankton Research,* 29**:** 483-493. DOI 10.1093/plankt/fbm031.

Bucklin, A. (2000). Population genetic variation of *Calanus finmarchicus* in Icelandic waters: preliminary evidence of genetic differences between Atlantic and Arctic populations. *ICES Journal of Marine Science,* 57**:** 1592-1604. DOI 10.1006/jmsc.2000.0957.

Bucklin, A., Kaartvedt, S., Guarnieri, M. & Goswami, U. (2000). Population genetics of drifting (*Calanus* spp.) and resident (*Acartia clausi*) plankton in Norwegian fjords. *Journal of Plankton Research,* 22**:** 1237-1251.

Bucklin, A., Sundt, R.C. & Dahle, G. (1996). The population genetics of *Calanus finmarchicus* in the North Atlantic. *Ophelia,* 44**:** 29-45.

Campbell, R.G., Wagner, M.M., Teegarden, G.J., Boudreau, C.A. & Durbin, E.G. (2001). Growth and development rates of the copepod *Calanus finmarchicus* reared in the laboratory. *Marine Ecology Progress Series,* 221**:** 161-183.

Carstensen, J. & Weydmann, A. (2012). Tipping points in the Arctic: eyeballing or statistical significance? *Ambio,* 41**:** 34-43.

Chust, G., Castellani, C., Licandro, P., Ibaibarriaga, L., Sagarminaga, Y. & Irigoien, X. (2013). Are *Calanus* spp. shifting poleward in the North Atlantic? A habitat modelling approach. *ICES Journal of Marine Science,* 71**:** 241-253. DOI 10.1093/icesjms/fst147.

Conover, R.J. & Siferd, T.D. (1993). Dark-season survival strategies of coastal zone zooplankton in the Canadian Arctic. *Arctic***:** 303-311.

Conover, R.J. (1988). Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. *Hydrobiologia,* 167**:** 127-142.

Crawford, D.L. & Oleksiak, M.F. (2016). Ecological population genomics in the marine environment. *Briefings in functional genomics,* 15**:** 342-351.

Cushing, D. (1990). Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in marine biology,* 26**:** 249-293.

Daase, M., Falk-Petersen, S., Varpe, Ø., Darnis, G., Søreide, J.E., Wold, A., Leu, E., Berge, J., Philippe, B. & Fortier, L. (2013). Timing of reproductive events in the marine copepod *Calanus glacialis*: a pan-Arctic perspective. *Canadian journal of fisheries and aquatic sciences,* 70**:** 871-884.

Daase, M. & Eiane, K. (2007). Mesozooplankton distribution in northern Svalbard waters in relation to hydrography. *Polar Biology,* 30**:** 969-981.

Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J.-M., Da Silva, C., Brinkmann, H., Mikhaleva, J., Olsen, L.C., Jubin, C. & Cañestro, C. (2010). Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science,* 330**:** 1381-1385.

Diel, S. & Tande, K. (1992). Does the spawning of *Calanus finmarchicus* in high latitudes follow a reproducible pattern? *Marine Biology,* 113**:** 21-31.

Falk-Petersen, S., Mayzaud, P., Kattner, G. & Sargent, J. (2009). Lipids and life strategy of Arctic *Calanus*. *Marine Biology Research,* 5**:** 18-39. DOI 10.1080/17451000802512267

Falk-Petersen, S., Pavlov, V., Timofeev, S. & Sargent, J.R. (2007). Climate variability and possible effects on arctic food chains: the role of *Calanus*. *Arctic alpine ecosystems and people in a changing environment Springer, Berlin***:** 147-166.

Falk-Petersen, S., Dahl, T.M., Scott, C.L., Sargent, J.R., Gulliksen, B., Kwasniewski, S., Hop, H. & Millar, R.-M. (2002). Lipid biomarkers and trophic linkages between ctenophores and copepods in Svalbard waters. *Marine Ecology Progress Series,* 227**:** 187-194.

Falk-Petersen, S., Pedersen, G., Kwasniewski, S., Hegseth, E.N. & Hop, H. (1999). Spatial distribution and life-cycle timing of zooplankton in the marginal ice zone of the Barents Sea during the summer melt season in 1995. *Journal of Plankton Research,* 21.

Fleminger, A. & Hulsemann, K. (1977). Geographical range and taxonomic divergence in North Atlantic *Calanus* (*C. helgolandicus*, *C. finmarchicus* and *C. glacialis*). *Marine Biology,* 40**:** 233-248.

Frost, B. (1989). A taxonomy of the marine calanoid copepod genus *Pseudocalanus. Canadian Journal of Zoology,* 67**:** 525-551.

Frost, B. (1974). *Calanus marshallae*, a new species of calanoid copepod closely allied to the sibling species *C. finmarchicus* and *C. glacialis*. *Marine Biology,* 26**:** 77-99.

Gabrielsen, T.M., Merkel, B., Søreide, J.E., Johansson-Karlsson, E., Bailey, A., Vogedes, D., Nygård, H., Varpe, Ø. & Berge, J. (2012). Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. *Polar Biology,* 35**:** 1621-1628. DOI 10.1007/s00300-012-1202-7.

Gatten, R., Corner, E., Kilvington, C. & Sargent, J. (2013). A Seasonal Survey of the Lipids in *Calanus helgolandicus* Claus from the English Channel. *Cyclic Phenomena in Marine Plants and Animals: Proceedings of the 13th European Marine Biology Symposium, Isle of Man, 27 September-4 October 1978.* Elsevier.

Gislason, A. & Astthorsson, O.S. (2002). The food of Norwegian spring-spawning herring in the western Norwegian Sea in relation to the annual cycle of zooplankton. *Sarsia: North Atlantic Marine Science,* 87**:** 236-247.

Grant, R.A. & Linse, K. (2009). Barcoding Antarctic Biodiversity: current status and the CAML initiative, a case study of marine invertebrates. *Polar Biology,* 32**:** 1629.

Hatlebakk, M.K. (2014). Capital or Income Breeder: The Role of Lipids and Fatty Acid Composition for Successful Reproduction in *Calanus glacialis*. Institutt for biologi.

Hays, G.C., Richardson, A.J. & Robinson, C. (2005). Climate change and marine plankton. *Trends in ecology & evolution,* 20**:** 337-344.

Helaouët, P. & Beaugrand, G. (2007). Macroecology of *Calanus finmarchicus* and *C. helgolandicus* in the North Atlantic Ocean and adjacent seas. *Marine Ecology Progress Series,* 345**:** 147-165.

Helyar, S.J., Hemmer‐Hansen, J., Bekkevold, D., Taylor, M., Ogden, R., Limborg, M., Cariani, A., Maes, G., Diopere, E. & Carvalho, G. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular ecology resources,* 11**:** 123-136.

Hirche, H.-J. & Kosobokova, K. (2007). Distribution of *Calanus finmarchicus* in the northern North Atlantic and Arctic Ocean—Expatriation and potential colonization. *Deep Sea Research Part II: Topical Studies in Oceanography,* 54**:** 2729-2747. DOI 10.1016/j.dsr2.2007.08.006.

Hirche, H.-J., Brey, T. & Niehoff, B. (2001). A high-frequency time series at Ocean Weather Ship Station M (Norwegian Sea): population dynamics of *Calanus finmarchicus*. *Marine Ecology Progress Series,* 219**:** 205-219.

Hirche, H.-J. (1997). Life cycle of the copepod *Calanus hyperboreus* in the Greenland Sea. *Marine Biology,* 128**:** 607-618.

Hirche, H.-J. & Niehoff, B. (1996). Reproduction of the Arctic copepod *Calanus hyperboreus* in the Greenland Sea-field and laboratory observations. *Polar biology,* 16**:** 209-219.

Hirche, H.-J. (1991). Distribution of dominant calanoid copepod species in the Greenland Sea during late fall. *Polar Biology,* 11**:** 351-362.

Hirche, H.-J. (1990). Egg production of *Calanus finmarchicus* at low temperature. *Marine Biology,* 106**:** 53-58.

Hirche, H.-J. (1983). Overwintering of *Calanus finmarchicus* and *Calanus helgolandicus*. *Marine ecology progress series Oldendorf,* 11**:** 281-290.

Hofmann, G.E., Burnaford, J.L. & Fielman, K.T. (2005). Genomics-fueled approaches to current challenges in marine ecology. *Trends in Ecology & Evolution,* 20**:** 305-311.

Huntley, M.E. & Lopez, M.D. (1992). Temperature-dependent production of marine copepods: a global synthesis. *The American Naturalist,* 140**:** 201-242.

Jansen, T. (2016). First‐year survival of North East Atlantic mackerel (*Scomber scombrus*) from 1998 to 2012 appears to be driven by availability of *Calanus*, a preferred copepod prey. *Fisheries Oceanography,* 25**:** 457-469.

Jaschnov, W. (1972). On the systematic status of *Calanus glacialis*, *Calanus finmarchicus* and *Calanus helgolandicus*. *Crustaceana,* 22**:** 279-284.

Jaschnov, W. (1970). Distribution of *Calanus* species in the seas of the northern hemisphere. *Internationale Revue der gesamten Hydrobiologie und Hydrographie,* 55**:** 197-212.

Jaschnov, W. (1966). Water mass and Plankton: 4. *Calanus finmarchicus***:** 404-412.

Jaschnov, V. (1961). Water masses and plankton. 1. Species of *Calanus finmarchicus* sl as indicators of definite water masses. *Zoologicheskii Zh,* 40**:** 1314-1334.

Jaschnov, W. (1957). Comparative morphology of the species *Calanus finmarchicus* s. 1. *Zool Zh,* 36**:** 191-198.

Jaschnov, W. (1955). Morphology, distribution and systematics of *Calanus finmarchicus* sl. *Zool Zh,* 34**:** 1210-1223.

Jaschnov, W. (1939). Life cycle and seasonal variability in distribution of age stages of *Calanus finmarchicus* in the Barents Sea. *Trudy VNIR,* 4**:** 225-44.

Kamshilov, M. (1955). Materials on the Biology of *Calanus finmarchicus* Gunner of the Barents and White Seas. *Tr Murmansk Biol St Akad Nauk SSSR,* 2**:** 62-86.

Kann, L. & Wishner, K. (1996). Genetic population structure of the copepod *Calanus finmarchicus* in the Gulf of Maine: allozyme and amplified mitochondrial DNA variation. *Marine Biology,* 125**:** 65-75.

Ki, J.-S., Hop, H., Kim, S.-J., Kim, I.-C., Park, H.G. & Lee, J.-S. (2010). Complete mitochondrial genome sequence of the Arctic gammarid, *Onisimus nanseni* (Crustacea; Amphipoda): Novel gene structures and unusual control region features. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics,* 5**:** 105-115.

Kim, S., Lim, B.J., Min, G.S. & Choi, H.G. (2013). The complete mitochondrial genome of Arctic *Calanus hyperboreus* (Copepoda, Calanoida) reveals characteristic patterns in calanoid mitochondrial genome. *Gene,* 520**:** 64-72. DOI 10.1016/j.gene.2012.09.059.

Kjellerup, S., Dünweber, M., Swalethorp, R., Nielsen, T.G., Møller, E.F., Markager, S. & Hansen, B.W. (2012). Effects of a future warmer ocean on the coexisting copepods *Calanus finmarchicus* and *C. glacialis* in Disko Bay, western Greenland. *Marine Ecology Progress Series,* 447**:** 87-108.

Kohn, A.B., Citarella, M.R., Kocot, K.M., Bobkova, Y.V., Halanych, K.M. & Moroz, L.L. (2012). Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, *Pleurobrachia bachei*. *Molecular phylogenetics and evolution,* 63**:** 203-207.

Kosobokova, K. (2012). Zooplankton of the Arctic Ocean: Community structure, Ecology, Spatial distribution. *GEOS, Moscow, 272p (in Russian)*.

Kosobokova, K.N., Hopcroft, R.R. & Hirche, H.-J. (2011). Patterns of zooplankton diversity through the depths of the Arctic's central basins. *Marine Biodiversity,* 41**:** 29-50.

Kosobokova, K. (1999). The reproductive cycle and life history of the Arctic copepod *Calanus glacialis* in the White Sea. *Polar Biology,* 22**:** 254-263.

Kosobokova, K., Hanssen, H., Hirche, H.-J. & Knickmeier, K. (1997). Composition and distribution of zooplankton in the Laptev Sea and adjacent Nansen Basin during summer, 1993. *Polar Biology,* 19**:** 63-76.

Kwasniewski, S., Hop, H., Falk-Petersen, S. & Pedersen, G. (2003). Distribution of *Calanus* species in Kongsfjorden, a glacial fjord in Svalbard. *Journal of plankton research,* 25**:** 1-20.

Lee, R.F., Hagen, W. & Kattner, G. (2006). Lipid storage in marine zooplankton. *Marine Ecology Progress Series,* 307**:** 273-306.

Leinaas, H.P., Jalal, M., Gabrielsen, T.M. & Hessen, D.O. (2016). Inter ‐ and intraspecific variation in body ‐ and genome size in calanoid copepods from temperate and arctic waters. *Ecology and evolution,* 6**:** 5585-5595.

Levinsen, H., Turner, J.T., Nielsen, T.G. & Hansen, B.W. (2000). On the trophic coupling between protists and copepods in arctic marine ecosystems. *Marine Ecology Progress Series,* 204**:** 65-77.

Lindeque, P.K., Hay, S.J., Heath, M.R., Ingvarsdottir, A., Rasmussen, J., Smerdon, G.R. & Waniek, J.J. (2006). Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. *Journal of Plankton Research,* 28**:** 221-238. DOI 10.1093/plankt/fbi115.

Lindeque, P., Harris, R., Jones, M. & Smerdon, G. (1999). Simple molecular method to distinguish the identity of *Calanus* species (Copepoda: Calanoida) at any developmental stage. *Marine Biology,* 133**:** 91-96.

Madoui, M.A., Poulain, J., Sugier, K., Wessner, M., Noel, B., Berline, L., Labadie, K., Cornils, A., Blanco‐Bercial, L. & Stemmann, L. (2017). New insights into global biogeography, population structure and natural selection from the genome of the epipelagic copepod *Oithona*. *Molecular Ecology*.

Madsen, S., Nielsen, T.G. & Hansen, B.W. (2001). Annual population development and production by *Calanus finmarchicus*, *C. glacialis* and *C. hyperboreus* in Disko Bay, western Greenland. *Marine Biology,* 139**:** 75-93.

Marlétaz, F., Le Parco, Y., Liu, S. & Peijnenburg, K.T.C.A. (2017). Extreme Mitogenomic Variation in Natural Populations of Chaetognaths. *Genome Biology and Evolution,* 9**:** 1374-1384. DOI 10.1093/gbe/evx090.

Matthews, J., Hestad, L. & Bakke, J. (1978). Ecological-studies in korsfjorden, western norway-generations and stocks of *Calanus hyperboreus* and *Calanus finmarchicus* in 1971-1974. *Oceanologica acta,* 1**:** 277-284.

Mclaren, I., Sevigny, J.-M. & Corkett, C. (1988). Body sizes, development rates, and genome sizes among *Calanus* species. *Biology of Copepods.* Springer.

Mcmanus, G.B. & Katz, L.A. (2009). Molecular and morphological methods for identifying plankton: what makes a successful marriage? *Journal of Plankton Research,* 31**:** 1119-1129.

Melle, W. & Skjoldal, H. (1998). Reproduction and development of *Calanus finmarchicus*, *C. glacialis* and *C. hyperboreus* in the Barents Sea. *Marine Ecology Progress Series,* 169**:** 211-228.

Michaud, J. & Taggart, C.T. (2007). Lipid and gross energy content of North Atlantic right whale food, *Calanus finmarchicus,* in the Bay of Fundy. *Endangered Species Research*.

Minxiao, W., Song, S., Chaolun, L. & Xin, S. (2011). Distinctive mitochondrial genome of Calanoid copepod *Calanus sinicus* with multiple large non-coding regions and reshuffled gene order: useful molecular markers for phylogenetic and population studies. *BMC Genomics,* 12**:** 73. DOI 10.1186/1471-2164-12-73.

Moroz, L.L., Kocot, K.M., Citarella, M.R., Dosung, S., Norekian, T.P., Povolotskaya, I.S., Grigorenko, A.P., Dailey, C., Berezikov, E. & Buckley, K.M. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature,* 510**:** 109.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R. & Hohenlohe, P.A. (2013). Genotyping‐by‐sequencing in ecological and conservation genomics. *Molecular ecology,* 22**:** 2841-2847.

Nelson, R.J., Carmack, E.C., Mclaughlin, F.A. & Cooper, G.A. (2009). Penetration of Pacific zooplankton into the western Arctic Ocean tracked with molecular population genetics. *Marine Ecology Progress Series,* 381**:** 129-138. DOI 10.3354/meps07940.

Niehoff, B. & Hirche, H.-J. (2005). Reproduction of *Calanus glacialis* in the Lurefjord (western Norway): indication for temperature-induced female dormancy. *Marine Ecology Progress Series,* 285**:** 107-115.

Nielsen, T.G., Kjellerup, S., Smolina, I., Hoarau, G. & Lindeque, P. (2014). Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phenological differences? *Marine Biology,* 161**:** 1299-1306. DOI 10.1007/s00227-014-2419-5.

Paffenhofer, G.A. (1976). Feeding, growth, and food conversion of the marine planktonic copepod *Calanus helgolandicus*. *Limnology and Oceanography,* 21**:** 39-50.

Paffenhöfer, G. & Strickland, J. (1970). A note on the feeding of *Calanus helgolandicus* on detritus. *Marine biology,* 5**:** 97-99.

Parent, G.J., Plourde, S. & Turgeon, J. (2012). Natural hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and Northwest Atlantic. *Limnology and Oceanography,* 57**:** 1057-1066. DOI 10.4319/lo.2012.57.4.1057.

Parent, G.J., Plourde, S. & Turgeon, J. (2011). Overlapping size ranges of *Calanus* spp. off the Canadian Arctic and Atlantic Coasts: impact on species' abundances. *Journal of Plankton Research,* 33**:** 1654-1665. DOI 10.1093/plankt/fbr072.

Pertsova, N. & Kosobokova, K. (2010). Interannual and seasonal variation of the population structure, abundance, and biomass of the arctic copepod *Calanus glacialis* in the White Sea. *Oceanology,* 50**:** 531-541.

Pett, W., Ryan, J.F., Pang, K., Mullikin, J.C., Martindale, M.Q., Baxevanis, A.D. & Lavrov, D.V. (2011). Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: insight from mtDNA and the nuclear genome. *Mitochondrial DNA,* 22**:** 130-142.

Planque, B. & Fromentin, J.-M. (1996). *Calanus* and environment in the eastern North Atlantic. I. Spatial and temporal patterns of *C. finmarchicus* and *C. helgolandicus*. *Marine Ecology Progress Series***:** 101-109.

Plourde, S., Joly, P., Runge, J.A., Zakardjian, B. & Dodson, J.J. (2001). Life cycle of C*alanus finmarchicus* in the lower St. Lawrence Estuary: the imprint of circulation and late timing of the spring phytoplankton bloom. *Canadian Journal of Fisheries and Aquatic Sciences,* 58**:** 647-658.

Poloczanska, E.S., Burrows, M.T., Brown, C.J., Garcia Molinos, J., Halpern, B.S., Hoegh-Guldberg, O., Kappel, C.V., Moore, P.J., Richardson, A.J. & Schoeman, D.S. (2016). Responses of marine organisms to climate change across oceans. *Frontiers in Marine Science,* 3**:** 62.

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A. & Savidge, G. (2009). High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus. Proceedings of the Royal Society of London B: Biological Sciences,* 276**:** 301-307.

Rees, C. (1957). *Continuous Plankton Records: the distribution of* Calanus finmarchicus *(Gunn.) in the North Sea and north-eastern Atlantic, 1946-1953*.

Rees, C.B. (1949). *Continuous Plankton Records: the Distribution of* Calanus Finmarchicus*, Gunn., and Its Two Forms in the North Sea, 1938-39*.

Reid, P.C., Edwards, M., Beaugrand, G., Skogen, M. & Stevens, D. (2003). Periodic changes in the zooplankton of the North Sea during the twentieth century linked to oceanic inflow. *Fisheries Oceanography,* 12**:** 260-269.

Reid, P.C., Borges, M.D. & Svendsen, E. (2001). A regime shift in the North Sea circa 1988 linked to changes in the North Sea horse mackerel fishery. *Fisheries Research,* 50**:** 163-171. DOI 10.1016/s0165-7836(00)00249-6.

Reitzel, A., Herrera, S., Layden, M., Martindale, M. & Shank, T. (2013). Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular ecology,* 22**:** 2953-2970.

Rey-Rassat, C., Irigoien, X., Harris, R., Head, R. & Carlotti, F. (2002). Egg production rates of *Calanus helgolandicus* females reared in the laboratory: variability due to present and past feeding conditions. *Marine Ecology Progress Series,* 238**:** 139-151.

Reygondeau, G. & Beaugrand, G. (2011). Future climate ‐ driven shifts in distribution of *Calanus finmarchicus*. *Global Change Biology,* 17**:** 756-766.

Richardson, A.J. (2008). In hot water: zooplankton and climate change. *Ices Journal of Marine Science,* 65**:** 279-295. DOI 10.1093/icesjms/fsn028.

Richardson, K., Jónasdóttir, S.H., Hay, S.J. & Christoffersen, A. (1999). *Calanus finmarchicus* egg production and food availability in the Faroe–Shetland Channel and northern North Sea: October–March. *Fisheries Oceanography,* 8**:** 153-162.

Riser, C.W., Wassmann, P., Reigstad, M. & Seuthe, L. (2008). Vertical flux regulation by zooplankton in the northern Barents Sea during Arctic spring. *Deep Sea Research Part II: Topical Studies in Oceanography,* 55**:** 2320-2329.

Ryan, J.F., Pang, K., Schnitzler, C.E., Nguyen, A.-D., Moreland, R.T., Simmons, D.K., Koch, B.J., Francis, W.R., Havlak, P. & Smith, S.A. (2013). The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science,* 342**:** 1242592.

Skreslet, S. & Rød, N.Å. (1986). Advection of *Calanus finmarchicus* between habitats in Norwegian coastal waters. *The Role of Freshwater Outflow in Coastal Marine Ecosystems.* Springer.

Slagstad, D., Ellingsen, I.H. & Wassmann, P. (2011). Evaluating primary and secondary production in an Arctic Ocean void of summer sea ice: An experimental simulation approach. *Progress in Oceanography,* 90**:** 117-131. DOI 10.1016/j.pocean.2011.02.009.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14**:** 1072-9. DOI 10.1111/1755-0998.12241.

Søreide, J.E., Leu, E., Berge, J., Graeve, M. & Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Global change biology,* 16**:** 3154-3163.

Søreide, J.E., Falk-Petersen, S., Hegseth, E.N., Hop, H., Carroll, M.L., Hobson, K.A. & Blachowiak-Samolyk, K. (2008). Seasonal feeding strategies of *Calanus* in the high-Arctic Svalbard region. *Deep Sea Research Part II: Topical Studies in Oceanography,* 55**:** 2225-2244.

Steen, H., Vogedes, D., Broms, F., Falk‐Petersen, S. & Berge, J. (2007). Little auks (*Alle alle*) breeding in a High Arctic fjord system: bimodal foraging strategies as a response to poor food quality? *Polar Research,* 26**:** 118-125.

Strugnell, J.M. & Lindgren, A.R. (2007). A barcode of life database for the Cephalopoda? Considerations and concerns. *Reviews in Fish Biology and Fisheries,* 17**:** 337-344.

Tande, K.S. (1982). Ecological investigations on the zooplankton community of Balsfjorden, northern Norway: generation cycles, and variations in body weight and body content of carbon and nitrogen related to overwintering and reproduction in the copepod *Calanus finmarchicus* (Gunnerus). *Journal of Experimental Marine Biology and Ecology,* 62**:** 129-142.

Trivedi, S., Aloufi, A.A., Ansari, A.A. & Ghosh, S.K. (2016). Role of DNA barcoding in marine biodiversity assessment and conservation: an update. *Saudi journal of biological sciences,* 23**:** 161-171.

Unal, E. & Bucklin, A. (2010). Basin-scale population genetic structure of the planktonic copepod *Calanus finmarchicus* in the North Atlantic Ocean. *Progress in Oceanography,* 87**:** 175-185. DOI 10.1016/j.pocean.2010.09.017.

Unstad, K.H. & Tande, K.S. (1991). Depth distribution of *Calanus finmarchicus* and *C. glacialis* in relation to environmental conditions in the Barents Sea. *Polar Research,* 10**:** 409-420.

Utne, K.R., Hjøllo, S.S., Huse, G. & Skogen, M. (2012). Estimating the consumption of *Calanus finmarchicus* by planktivorous fish in the Norwegian Sea using a fully coupled 3D model system. *Marine Biology Research,* 8**:** 527-547.

Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A. & Irigoien, X. (2015). Modelling the future biogeography of North Atlantic zooplankton communities in response to climate change. *Marine Ecology Progress Series,* 531**:** 121-142. DOI 10.3354/meps11299.

Vogedes, D., Varpe, Ø., Søreide, J.E., Graeve, M., Berge, J. & Falk-Petersen, S. (2010). Lipid sac area as a proxy for individual lipid content of arctic calanoid copepods. *Journal of plankton research,* 32**:** 1471-1477.

Wassmann, P., Kosobokova, K., Slagstad, D., Drinkwater, K., Hopcroft, R., Moore, S., Ellingsen, I., Nelson, R., Carmack, E. & Popova, E. (2015). The contiguous domains of Arctic Ocean advection: trails of life and death. *Progress in Oceanography,* 139**:** 42-65.

Wassmann, P., Duarte, C.M., Agusti, S. & Sejr, M.K. (2011). Footprints of climate change in the Arctic marine ecosystem. *Global change biology,* 17**:** 1235-1249.

Wassmann, P., Reigstad, M., Haug, T., Rudels, B., Carroll, M.L., Hop, H., Gabrielsen, G.W., Falk-Petersen, S., Denisenko, S.G. & Arashkevich, E. (2006). Food webs and carbon flux in the Barents Sea. *Progress in Oceanography,* 71**:** 232-287.

Węsławski, J., Stempniewicz, L., Mehlum, F. & Kwaśniewski, S. (1999). Summer feeding strategy of the little auk (*Alle alle*) from Bjørnøya, Barents Sea. *Polar Biology,* 21**:** 129-134.

Weydmann, A., Coelho, N.C., Serrão, E.A., Burzyński, A. & Pearson, G.A. (2016). Pan-Arctic population of the keystone copepod *Calanus glacialis*. *Polar Biology***:** 1-8. DOI 10.1007/s00300-016-1898-x.

Wiebe, P.H., Bucklin, A., Madin, L., Angel, M.V., Sutton, T., Pagés, F., Hopcroft, R.R. & Lindsay, D. (2010). Deep-sea sampling on CMarZ cruises in the Atlantic Ocean–an Introduction. *Deep Sea Research Part II: Topical Studies in Oceanography,* 57**:** 2157-2166.

Williams, R. (1985). Vertical distribution of *Calanus finmarchicus* and *C. helgolandicus* in relation to the development of the seasonal thermocline in the Celtic Sea. *Marine Biology,* 86**:** 145-149.

Wilson, S.E., Steinberg, D.K. & Buesseler, K.O. (2008). Changes in fecal pellet characteristics with depth as indicators of zooplankton repackaging of particles in the mesopelagic zone of the subtropical and subarctic North Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography,* 55**:** 1636-1647.

Yebra, L., Bonnet, D., Harris, R.P., Lindeque, P.K. & Peijnenburg, K.T. (2011). Barriers in the pelagic: population structuring of *Calanus helgolandicus* and *C. euxinus* in European waters. *Marine Ecology Progress Series,* 428**:** 135-149.

Paper I

Genetics redraws pelagic biogeography of *Calanus*

Marvin Choquet[1*], Maja Hatlebakk[1,2], Anusha KS Dhanasiri[1], Ksenia Kosobokova[3], Irina Smolina[1], Janne E Søreide[2], Camilla Svensen[4], Webjørn Melle[5], Sławomir Kwaśniewski[6], Ketil Eiane[1], Malin Daase[4], Vigdis Tverberg[1], Stig Skreslet[1], Ann Bucklin[7], Galice Hoarau[1]

[1]*Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway*
[2]*Department of Arctic Biology, The University Centre in Svalbard, Longyearbyen, Norway*
[3]*Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia*
[4]*Department of Arctic and Marine Biology, UiT-the Arctic University of Norway, Tromsø, Norway*
[5]*Institute of Marine Research, Bergen, Norway*
[6]*Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland*
[7]*University of Connecticut, Groton, USA*

Abstract

Planktonic copepods of the genus *Calanus* play a central role in North Atlantic/Arctic marine food webs. Here, using molecular markers, we redrew the distributional ranges of *Calanus* species inhabiting the North Atlantic and Arctic Oceans and revealed much wider and overlapping distributions than previously described. The Arctic shelf species, *C. glacialis,* dominated the zooplankton assemblage of many Norwegian fjords, where only *C. finmarchicus* has been reported previously. In these fjords, high occurrences of the Arctic species *C. hyperboreus* were also found. Molecular markers revealed that the most common method of species identification, prosome length, cannot reliably discriminate the species in Norwegian fjords. Differences in degree of genetic differentiation among fjord populations of the two species suggested that *C. glacialis* is a more permanent resident of the fjords than *C. finmarchicus*. We found no evidence of hybridisation between the species. Our results indicate a critical need for the wider use of molecular markers to reliably identify and discriminate these morphologically-similar copepod species, which serve as important indicators of climate responses.

1.Introduction

Copepods of the genus *Calanus* are central in North Atlantic and Arctic pelagic food webs. Rich in lipids, they form a key link between primary producers and secondary consumers and predators. Four species of the genus *Calanus* occur throughout the North Atlantic and Arctic Oceans (Fig.1): *C. helgolandicus (Chel), C. hyperboreus (Chyp), C. finmarchicus (Cfin)* and *C. glacialis (Cgla);* and there has been considerable effort to document and model their distributional changes (Beaugrand et al., 2002, Villarino et al., 2015). Importantly, abundances and dynamics of fish stocks are strongly associated with *Calanus* species composition and abundances (Beaugrand and Kirby, 2010), and climate-driven changes in their biogeographical distributions (i.e., range shifts) can lead to ecosystem regime shifts and potential collapse of fish stocks such as cod (Beaugrand et al., 2003). However, distinguishing *Calanus* species is challenging, due to their morphological similarity and lack of diagnostic characters. The usual method of species identification is body (prosome) length, although this approach has been questioned (Gabrielsen et al., 2012). Misidentification may thus occur, impacting the reliability of our current knowledge of species distributions, and preventing accurate assessment of species geographic range shifts in response to climate change.

Here we re-examine the distributional ranges of four co-occurring *Calanus* species in the North Atlantic and Arctic Oceans, using six molecular markers designed to ensure reliable species identification.

2.Material and Methods

Sample collection

Zooplankton samples were collected from 83 locations in the North Atlantic and Arctic Oceans (Supplementary1) by vertical nets tows with 150-200μm mesh sizes and preserved in 70-80% ethanol. A Folsom plankton splitter was used to make subsamples containing up to 150 *Calanus* individuals from developmental stage CIV to CVI (Supplementary1). No morphological identification was performed for any individuals.

Molecular species identification

DNA was extracted from the excised antennae of each specimen, using the HotSHOT protocol (Montero-Pau et al., 2008) and molecular species identification of 4,434 individuals was achieved using six nuclear markers type InDels (Insertion or Deletion motifs) (Smolina et al., 2014) scored on a 3500xL Genetic Analyzer (Applied Biosystems). These bi-parentally inherited markers are easy to use and can potentially

detect hybridization (Nielsen et al., 2014). Their reliability was confirmed by the traditional, but more cost and labor intensive mitochondrial 16S rDNA sequencing (mtDNA) (Lindeque et al., 2006, Lindeque et al., 1999) of 159 individuals from 53 locations (Supplementary2-3), following Smolina *et al*. (2014). In addition, 129 individuals from Saltfjord/Skjerstadfjord were measured (prosome length) and sequenced for the 16S (Table 1; Supplementary4-5). Identification of specimens from InDels and 16S rDNA sequences was congruent for all 677 individuals investigated (288 in present study – Supplementary2-3-4; and 389 in Nielsen *et al*., 2014). InDel markers were also used to test for the presence of hybrids between *Cfin* and *Cgla* (Nielsen et al., 2014) (Supplementary6).

Population differentiation

Population genetic analysis was carried out to distinguish between fjord resident and drifting species (Bucklin et al., 2000) (Supplementary7). Focusing on *Cfin* and *Cgla* populations, genetic differentiation was measured using the global index of population differentiation, F_{ST} (Weir and Cockerham, 1984), based on 10 microsatellite DNA markers (Parent et al., 2012, Provan et al., 2009) assayed for 24 individuals per species from 3 locations: Isfjord, Saltfjord and Lurefjord.

3.Results and Discussion

Identification of *Calanus* species using molecular markers revealed much wider distributional ranges than previously reported (Figs.1&2, Supplementary1). The distribution of *Chel* was known to extend from the Mediterranean Sea to the North Sea (58°N – Fig.1) (Barnard et al., 2004). Here, we identified *Chel* in several Norwegian fjords and in the Norwegian Sea as far north as 70° N (Fig.2). Specimens found in Myken stations (66°N) and near Tromsø (70°N) could result from transport in ocean frontal jet currents running from the North Sea along the Norwegian coast. However, the high prevalence (85%) of the species recorded in the relatively isolated Sognefjord (61°N) may represent a locally-established population. It remains to be tested whether *Chel* has always been present in these fjords but never identified, or whether our findings represent evidence of a recent biogeographical range shift.

Previous reports of the Arctic *Chyp* (Conover, 1988) occurring in the northern Norwegian Sea (Fig.1) have been attributed to transport of individuals by Arctic intermediate waters (Broms et al., 2009). Here, we detected the species in large proportions along the Norwegian coast, everywhere north of 58°N (Fig.2,

Supplementary1). Whether the southern presence of *Chyp* results from advection from Arctic stocks or from self-reproducing populations remains to be investigated.

Calanus finmarchicus is currently considered to be an indicator species of North Atlantic water masses (Conover, 1988), and our results largely support this view (Fig.2). The genetically confirmed species range extends as far north as 87°N and as far east in the Arctic as the eastern boarder of the Laptev Sea (78°N, 113°E – Fig.2), regions of the Arctic Ocean affected by Atlantic inflow. It was proposed that *Cfin* may thrive in these Northern regions and replace *Cgla* in response to Arctic warming (Wassmann et al., 2015), however, at present the individuals recorded at these most northerly locations were likely transported from southern populations (Wassmann et al., 2015).

Calanus glacialis is regarded as a true Arctic shelf species, which serves as a circumpolar indicator of these waters (Conover, 1988) (Fig.1). We rarely observed it offshore in Atlantic waters, but documented the species occurrence in many Norwegian fjords, as far south as 60°N (Fig.2), where it usually co-occurred with *Cfin* in fjords with deep basins separated from shelf waters by shallower sills (Supplementary1). In several fjords, *Cgla* dominated over other *Calanus* species; we recorded a positive gradient of relative abundance of *Cgla* from the mouth to the innermost areas of some fjords (e.g. Ranfjord, Supplementary1).

In the fjords, prosome length of *Cgla* and *Cfin* overlapped completely (Table 1; Supplementary5), which explains why *Cgla*'s large occurrence has not been reported previously. Furthermore, a recent study has concluded that morphological characters cannot reliably distinguish between *Cfin* and *Cgla* throughout their range (Choquet et al., unpubl.).

Some zooplankton species are long-term residents of Norwegian fjords, while others are replaced periodically with basin water exchanges (Lindahl and Hernroth, 1988). Resident species are expected to show greater genetic differentiation among fjord populations than drifting species (Bucklin et al., 2000). Our analysis found no significant genetic differentiation among fjord populations of *Cfin* (F_{ST}=0.004^{NS}), but *Cgla* populations did differ significantly (F_{ST}=0.03*), suggesting lower rates of exchange (i.e., gene flow) for *Cgla* than for *Cfin.* These results support previous descriptions of *Cfin* as a drifting species (Bucklin et al., 2000) that is advected into and out of fjords seasonally (Skreslet and Rød, 1986). Less gene flow – together with the absence of offshore populations – suggests that *Cgla* populations are resident (Bucklin et al., 2000). In both the White Sea (Pertsova and Kosobokova, 2010) and Lurefjord (Niehoff and Hirche, 2005), *Cgla* is known to migrate in early summer from warm surface layers to colder deep water. This may explain the species ability to maintain local populations and avoid transport out of fjords.

Hybridisation between *Cfin* and *Cgla* has been suggested in the Northwest Atlantic (Parent et al., 2012). Notably, no first-generation hybrids were found in our

survey of 4,434 individuals from samples collected throughout the Northeast Atlantic and Arctic Oceans (Supplementary6). Based on the nature of the molecular characters (InDels) used for species identification and careful ground-truthing of our molecular results, we conclude that hybridisation between the species, if it occurs at all, is rare or episodic.

Conclusion

Marine zooplankton have been regarded as sentinels of climate change (Hays et al., 2005), due to their short life histories and rapid responses to environmental variation. Development and use of molecular characters that can ensure accurate and reliable identification and discrimination of key indicator species, such as *Calanus*, is critically needed. Only then can these species be used to document past, present and future patterns of biogeographical distributions, and detect and track responses of pelagic communities to climate change.

Acknowledgements:

We thank M. Krogstad, E. Abramova, F. Norrbin, Ø. Leiknes, S. Basedow, T. Dale, T. Falkenhaug, A. Mailli, K. Last, S. Wells and the captains and crews of R/V Helmer Hanssen and G.O. Sars for their assistance with sampling. We are grateful to the ARCTOS network, for support and useful discussions. We acknowledge two anonymous reviewers for constructive comments.

Data accessibility:

Protocols are attached as Electronic Supplementary Material; genotypes and sequences have been deposited to public database, respectively Dryad (XXXX) and Genbank® (MF959702 to MF959730, and MF972920 to MF972922).

Funding:

MC was supported by The European Commission (FP7-EURO-BASIN-264933), the Norwegian Research Council (216578; 227139; 246747) and Nord University. MH was supported by UNIS. KK was supported by the Russian Foundation for Basic Research (15-29-02447; 16-04-00375) and the Russian Scientific Foundation (14-50-00095). MD was supported by NRC-226417. SK was supported by the Polish-Norwegian Research Program (Pol-Nor/201992/93/2014).

Competing interests:

We have no competing interests.

Ethical statement:

Ethical assessment not required.

References:

Barnard, R., Batten, S., Beaugrand, G., Buckland, C., Conway, D., Edwards, M., Finlayson, J., Gregory, L., Halliday, N. & John, A. (2004). Continuous plankton records: Plankton atlas of the North Atlantic Ocean (1958-1999). II. Biogeographical charts. *Marine Ecology-Progress Series*: 11-75.

Beaugrand, G. & Kirby, R.R. (2010). Climate, plankton and cod. *Global Change Biology,* 16: 1268-1280. DOI 10.1111/j.1365-2486.2009.02063.x.

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S. & Reid, P.C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature,* 426: 661-664.

Beaugrand, G., Reid, P.C., Ibanez, F., Lindley, J.A. & Edwards, M. (2002). Reorganization of North Atlantic marine copepod biodiversity and climate. *Science,* 296: 1692-1694.

Broms, C., Melle, W. & Kaartvedt, S. (2009). Oceanic distribution and life cycle of *Calanus* species in the Norwegian Sea and adjacent waters. *Deep Sea Research Part II: Topical Studies in Oceanography,* 56: 1910-1921. DOI 10.1016/j.dsr2.2008.11.005.

Bucklin, A., Kaartvedt, S., Guarnieri, M. & Goswami, U. (2000). Population genetics of drifting (*Calanus* spp.) and resident (*Acartia clausi*) plankton in Norwegian fjords. *Journal of Plankton Research,* 22: 1237-1251.

Conover, R. (1988). Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. *Biology of Copepods.* Springer.

Gabrielsen, T.M., Merkel, B., Søreide, J.E., Johansson-Karlsson, E., Bailey, A., Vogedes, D., Nygård, H., Varpe, Ø. & Berge, J. (2012). Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. *Polar Biology,* 35: 1621-1628. DOI 10.1007/s00300-012-1202-7.

Hays, G.C., Richardson, A.J. & Robinson, C. (2005). Climate change and marine plankton. *Trends in ecology & evolution,* 20: 337-344.

Lindahl, O. & Hernroth, L. (1988). Large-scale and long-term variations in the zooplankton community of the Gullmar fjord, Sweden, in relation to advective processes. *Marine Ecology Progress Series*: 161-171.

Lindeque, P.K., Hay, S.J., Heath, M.R., Ingvarsdottir, A., Rasmussen, J., Smerdon, G.R. & Waniek, J.J. (2006). Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. *Journal of Plankton Research,* 28: 221-238. DOI 10.1093/plankt/fbi115.

Lindeque, P., Harris, R., Jones, M. & Smerdon, G. (1999). Simple molecular method to distinguish the identity of *Calanus* species (Copepoda: Calanoida) at any developmental stage. *Marine Biology,* 133: 91-96.

Montero‐Pau, J., Gómez, A. & Muñoz, J. (2008). Application of an inexpensive and high‐throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods,* 6: 218-222.

Niehoff, B. & Hirche, H.-J. (2005). Reproduction of *Calanus glacialis* in the Lurefjord (western Norway): indication for temperature-induced female dormancy. *Marine Ecology Progress Series,* 285: 107-115.

Nielsen, T.G., Kjellerup, S., Smolina, I., Hoarau, G. & Lindeque, P. (2014). Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phenological differences? *Marine Biology,* 161**:** 1299-1306. DOI 10.1007/s00227-014-2419-5.

Parent, G.J., Plourde, S. & Turgeon, J. (2012). Natural hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and Northwest Atlantic. *Limnology and Oceanography,* 57**:** 1057-1066. DOI 10.4319/lo.2012.57.4.1057.

Pertsova, N. & Kosobokova, K. (2010). Interannual and seasonal variation of the population structure, abundance, and biomass of the arctic copepod *Calanus glacialis* in the White Sea. *Oceanology,* 50**:** 531-541.

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A. & Savidge, G. (2009). High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus*. *Proceedings Biological sciences / The Royal Society,* 276**:** 301-7. DOI 10.1098/rspb.2008.1062.

Skreslet, S. & Rød, N.Å. (1986). Advection of *Calanus finmarchicus* between habitats in Norwegian coastal waters. *The Role of Freshwater Outflow in Coastal Marine Ecosystems.* Springer.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14**:** 1072-9. DOI 10.1111/1755-0998.12241.

Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A. & Irigoien, X. (2015). Modelling the future biogeography of North Atlantic zooplankton communities in response to climate change. *Marine Ecology Progress Series,* 531**:** 121-142. DOI 10.3354/meps11299.

Wassmann, P., Kosobokova, K., Slagstad, D., Drinkwater, K., Hopcroft, R., Moore, S., Ellingsen, I., Nelson, R., Carmack, E. & Popova, E. (2015). The contiguous domains of Arctic Ocean advection: trails of life and death. *Progress in Oceanography,* 139**:** 42-65.

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

Tables

Table1: Comparison of *Calanus finmarchicus* (*Cfin*) and *C. glacialis* (*Cgla*) identification methods in Saltenfjord/Skjerstadfjord.

Saltenfjord / Skjerstadfjord	InDel Species ID	16S rDNA Species ID	Markers congruence	Prosome Length Range (µm)			
				N	Stage CV	N	Stage CVI female
Cfin	89	89	100%	26	1976.64 – 2717.76	14	2406.89 – 2747.02
Cgla	40	40	100%	20	2119.40 – 2623.33	69	2150.68 – 3030.50

Figures Legends

Fig.1: *Calanus* species distributional ranges in the North Atlantic and Arctic Oceans based on morphological identification.

For each panel, dark-shaded colour represents core area for each species, where reproduction is known to occur; light-shaded colour represents the total described distributional area (c.f. Supplementary8).

Fig.2: *Calanus* species distributional ranges in the North Atlantic and Arctic Oceans based on molecular species identifications. Pie charts represent relative frequencies of *C. glacialis* (blue), *C. finmarchicus* (red), *C. hyperboreus* (green) and *C. helgolandicus* (yellow) in each sample. * Indicates non-quantitative species records.

Figures

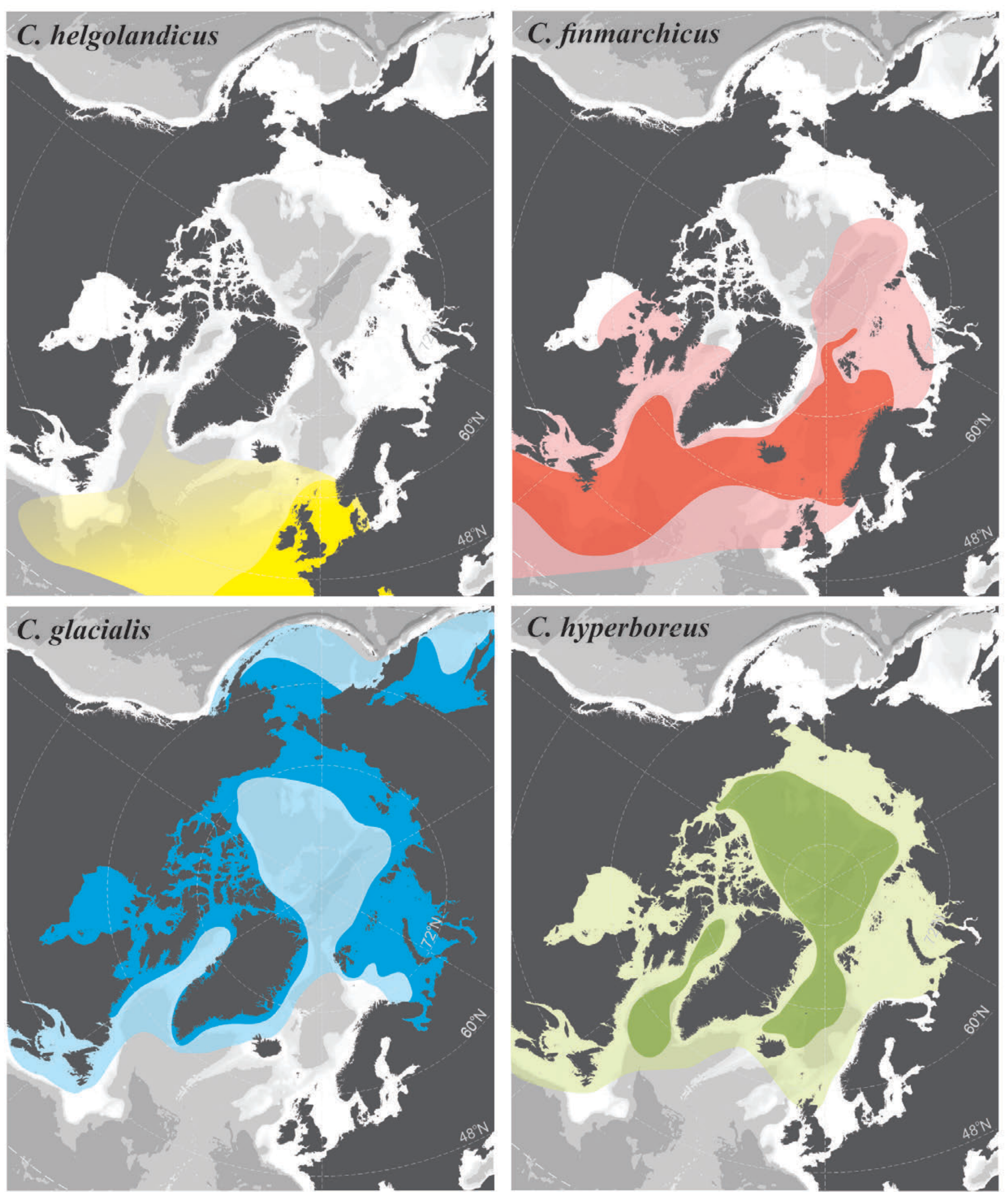


Fig. 1.

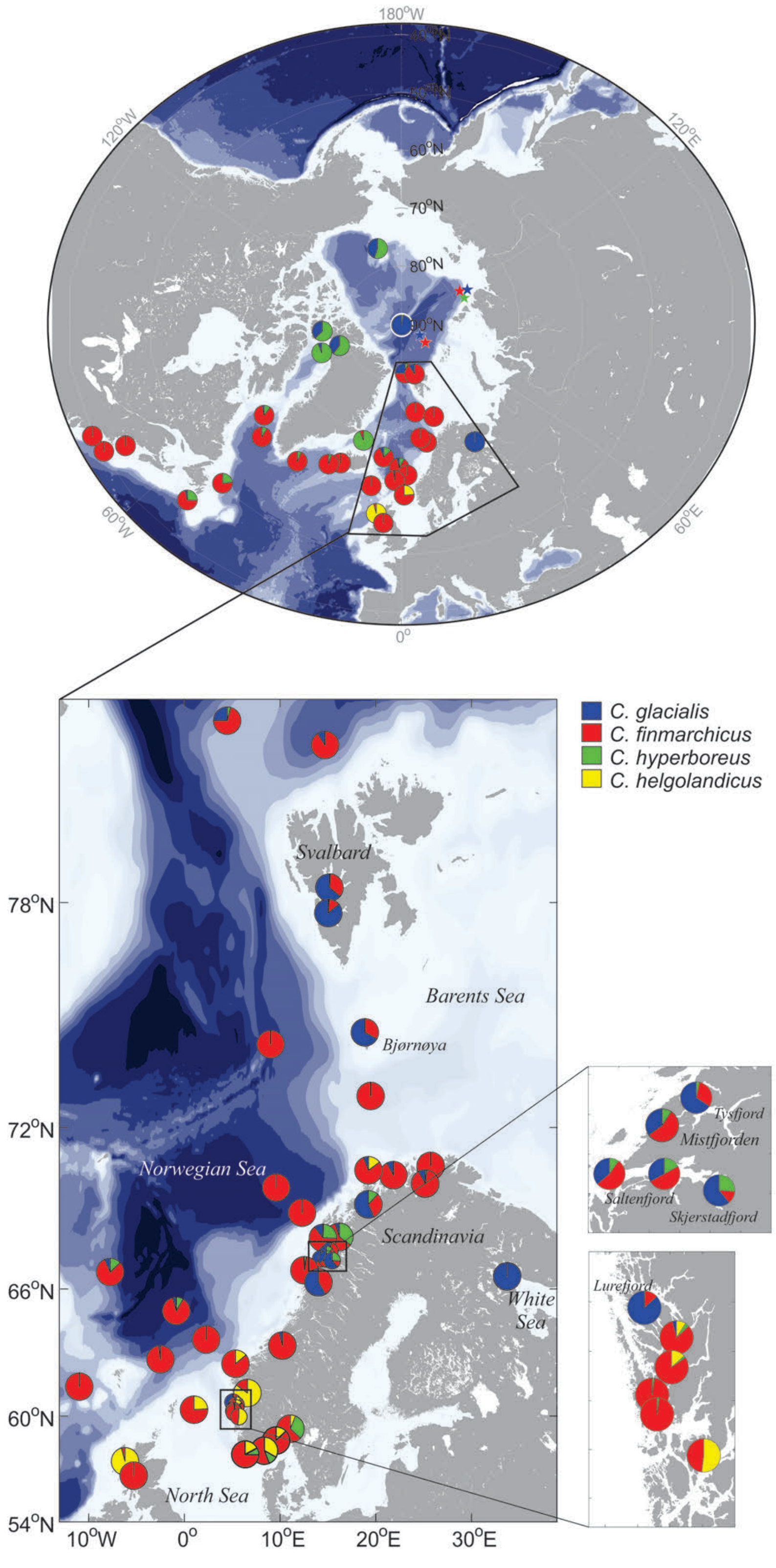


Fig. 2

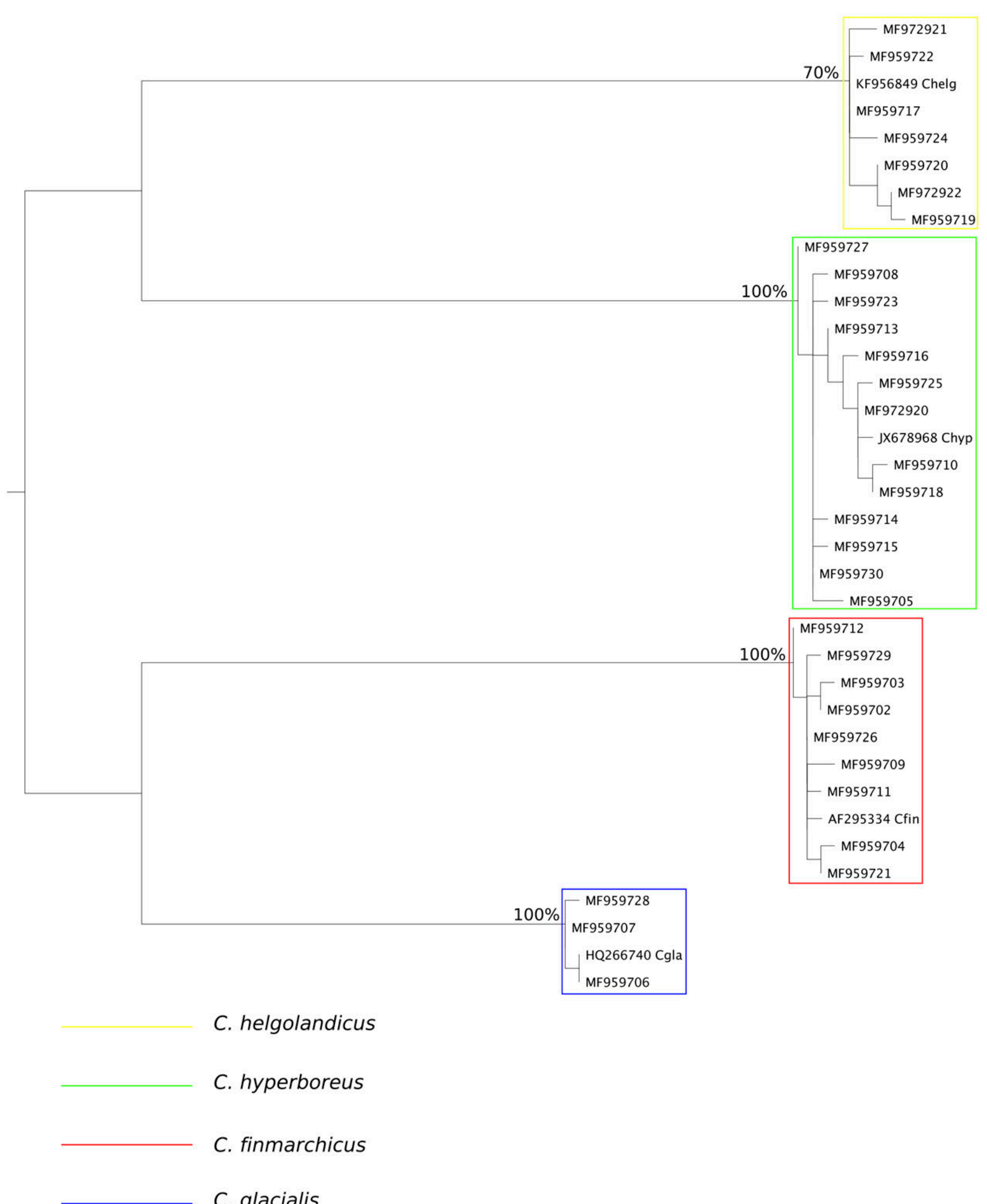
Electronic Supplementary Material

SUPPLEMENTARY 1

Sampling information and molecular-based *Calanus* species composition within samples. N individuals per sample were genotyped with nuclear InDel markers. Proportions of *C. glacialis* (*Cgla*), *C. finmarchicus* (*Cfin*), *C. hyperboreus* (*Chyp*) and *C. helgolandicus* (*Chel*) are reported. In the Laptev Sea and the Nansen Basin, only species presence (x) or absence (-) is reported.

Locations	GPS	Date	Depth	N	%*Cgla*	%*Cfin*	%*Chyp*	%*Chel*	Collector
North Pole	89°43 N 14°11 E	4/9/2012	100-0 m	48	98%	0%	2%	0%	J. Søreide
Nansen Basin	87° N 55°47 E	10/4/2016	4400-0 m	94	x	x	-	-	K. Kosobokova
Ice West Svalbard	81°22 N 4°27 E	10/1/2012	200-0 m	48	25%	71%	4%	0%	I. Smolina
Ice North Svalbard	80°58 N 14°41 E	1/15/2016	200-0 m	44	9%	91%	0%	0%	M. Choquet
West Greenland 1	78°42 N 70°43 W	9/6/2013	100-0 m	48	40%	0%	60%	0%	J. Søreide
Laptev Sea	78°30 N 113°01 E	9/6/2013	100-0 m	91	x	x	x	-	E. Abramova
Isfjord	78°19 N 15°09 E	6/5/2016	20-0 m	138	64%	35%	1%	0%	M. Hatlebakk
West Greenland 2	77°51 N 71°49 W	9/9/2013	95-0 m	48	6%	0%	94%	0%	J. Søreide
Van Mijenfjord	77°46 N 15°02 E	6/3/2016	20-0 m	91	86%	13%	1%	0%	M. Hatlebakk
Chukchi Sea	76°24 N 162°14 W	7/27/2016	300-0 m	60	45%	0%	55%	0%	A. Bucklin
West Greenland 3	76°23 N 84°30 W	9/15/2013	100-0 m	48	38%	0%	63%	0%	J. Søreide
Bjørnøya 1	74°50 N 18°51 E	08/08/2012	50-0 m	48	67%	33%	0%	0%	I. Smolina
Bjørnøya 2	74°30 N 9°01 E	8/9/2013	200-0 m	47	0%	100%	0%	0%	I. Smolina
Fugløya-Bjørnøya	72°59 N 19°27 E	8/7/2013	100-0 m	47	0%	100%	0%	0%	I. Smolina
Porsangerfjord 1	70°43 N 25°44 E	8/21/2013	150-0 m	46	0%	100%	0%	0%	F. Norrbin
Porsangerfjord 2	70°40 N 25°39 E	6/3/2016	120-0 m	48	0%	100%	0%	0%	C. Svensen
Off-Tromsø	70°34 N 19°14 E	1/8/2016	300-0 m	79	6%	79%	0%	15%	M. Choquet
Lopphavet	70°25 N 21°50 E	6/2/2016	200-0 m	48	8%	92%	0%	0%	F. Norrbin
Østerbotn	70°06 N 25°09 E	8/25/2013	100-0 m	39	8%	90%	2%	0%	F. Norrbin
Gimsøy 1	69°57 N 9°35 E	8/5/2013	200-0 m	47	0%	100%	0%	0%	I. Smolina
Balsfjord 1	69°21 N 19°13 E	5/31/2016	170-0 m	39	90%	8%	2%	0%	C. Svensen
Balsfjord 2	69°21 N 19°13 E	5/3/2016	170-0 m	46	28%	52%	20%	0%	C. Svensen
Gimsøy 2	69°2 N 12°17 E	8/5/2012	200-0 m	46	0%	100%	0%	0%	I. Smolina
North Iceland	68°48 N 18°23 W	5/12/2013	200-0 m	48	0%	6%	94%	0%	I. Smolina
Tysfjord	68°07 N 16°11 E	4/9/2016	200-0 m	37	5%	60%	35%	0%	F. Norrbin
Vestfjord 1	68°06 N 14°28 E	5/30/2016	270-0 m	49	12%	76%	10%	2%	M. Krogstad
Vestfjord 2	68°05 N 14°58 E	4/10/2016	200-0 m	43	12%	21%	67%	0%	F. Norrbin
Vestfjord 3	68°02 N 14°27 E	6/20/2016	290-0 m	48	8%	90%	2%	0%	M. Krogstad
Sørfoldfjord 1	67°35 N 14°50 E	4/20/2016	510-0 m	43	51%	42%	7%	0%	M. Krogstad
Sørfoldfjord 2	67°34 N 15°11 E	4/20/2016	460-0 m	48	67%	29%	4%	0%	M. Krogstad
Sørfoldfjord 3	67°30 N 15°28 E	4/20/2016	480-0 m	55	76%	24%	0%	0%	M. Krogstad
Mistfjord 1	67°27 N 14°51 E	10/29/2013	290-0 m	40	25%	70%	5%	0%	M. Krogstad
Mistfjord 2	67°27 N 14°50 E	2/17/2016	285-0 m	82	45%	42%	13%	0%	M. Krogstad
Saltfjord 1	67°16 N 14°38 E	2/15/2016	375-0 m	98	42%	40%	18%	0%	M. Krogstad
Saltfjord 2	67°16 N 14°38 E	7/6/2016	360-0 m	93	30%	69%	1%	0%	M. Krogstad
Skjerstadfjord 1	67°15 N 14°50 E	7/12/2016	500-0 m	89	33%	50%	17%	0%	M. Krogstad
Skjerstadfjord 2	67°11 N 15°26 E	2/26/2016	385-0 m	151	61%	13%	26%	0%	M. Krogstad
Myken 1	67°03 N 13°31 E	8/3/2016	200-0 m	48	2%	96%	2%	0%	M. Krogstad
Myken 2	66°48 N 12°32 E	8/3/2016	210-0 m	45	0%	100%	0%	0%	M. Krogstad
Myken 3	66°46 N 12°30 E	8/3/2016	95-0 m	12	0%	100%	0%	0%	M. Krogstad
Myken 4	66°46 N 12°28 E	8/3/2016	28-0 m	11	0%	73%	0%	27%	M. Krogstad
Myken 5	66°46 N 12°23 E	8/3/2016	140-0 m	12	0%	100%	0%	0%	M. Krogstad
Myken 6	66°45 N 12°28 E	8/3/2016	95-0 m	12	0%	100%	0%	0%	M. Krogstad
North Atlantic 1	66°43 N 7°46 W	5/8/2013	1000-0 m	46	7%	80%	13%	0%	I. Smolina
White Sea	66°33 N 33°43 E	8/22/2016	100-0 m	115	100%	0%	0%	0%	K. Kosobokova
Ranfjord 1	66°17 N 14°00 E	6/14/2016	410-0 m	47	98%	2%	0%	0%	M. Krogstad
Ranfjord 2	66°14 N 13°24 E	6/14/2016	275-0 m	46	63%	37%	0%	0%	M. Krogstad
Ranfjord 3	66°12 N 12°49 E	6/14/2016	215-0 m	45	13%	87%	0%	0%	M. Krogstad
North Atlantic 2	65°03 N 00°51 W	5/5/2013	1000-0 m	48	4%	88%	8%	0%	I. Smolina
South Iceland	63°49 N 24°18 W	5/14/2013	200-0 m	48	0%	100%	0%	0%	I. Smolina
Steinkjer	63°49 N 11°18 E	6/14/2016	50-0 m	45	0%	98%	0%	2%	G. Hoarau & K. Eiane
North Atlantic 3	63°45 N 02°16 E	5/4/2013	200-0 m	47	0%	100%	0%	0%	I. Smolina
Trondheimsfjord	63°29 N 10°14 E	8/22/2016	450-0 m	48	4%	96%	0%	0%	Ø. Leiknes
North Atlantic 4	62°50 N 28°17 W	5/15/2013	1000-0 m	40	0%	95%	5%	0%	I. Smolina
North Atlantic 5	62°50 N 02°31 W	4/13/2012	150-0 m	48	2%	98%	0%	0%	S. Basedow
Svinøy	62°37 N 05°20 E	01/08/12	127-0 m	51	0%	86%	0%	14%	I. Smolina
Labrador 1	62°13 N 57°21 W	5/26/2013	1000-0 m	95	3%	86%	11%	0%	I. Smolina
North Atlantic 6	61°30 N 10°59 W	4/10/2012	50-0 m	48	0%	100%	0%	0%	S. Basedow
Sognefjord	61°11 N 06°35 E	6/22/2016	200-0 m	41	0%	15%	0%	85%	T. Dale
Lurefjord 1	60°43 N 05°04 E	6/22/2016	230-0 m	48	81%	19%	0%	0%	M. Choquet
Lurefjord 2	60°41 N 05°09 E	6/22/2016	340-0 m	92	88%	12%	0%	0%	M. Choquet
Lurefjord 3	60°37 N 05°11 E	6/22/2016	160-0 m	49	90%	8%	0%	2%	M. Choquet
Christianssund	60°36 N 38°20 W	6/1/2013	1000-0 m	45	0%	93%	7%	0%	I. Smolina
Osterfjord	60°34 N 05°24 E	6/23/2016	550-0 m	47	4%	83%	4%	9%	M. Choquet
Sørfjord 1	60°31 N 05°21 E	6/23/2016	487-0 m	48	0%	94%	6%	0%	M. Choquet
Sørfjord 2	60°28 N 05°40 E	6/23/2016	360-0 m	43	0%	81%	0%	19%	M. Choquet
Sørfjord 3	60°26 N 05°30 E	6/23/2016	225-0 m	48	0%	79%	0%	21%	M. Choquet
North Sea	60°20 N 01°00 E	4/2/2012	50-0 m	45	0%	76%	0%	24%	S. Basedow
Raunefjord	60°17 N 05°08 E	6/4/2016	180-0 m	45	0%	98%	2%	0%	W. Melle
Korsfjord	60°11 N 05°12 E	6/6/2016	200-0 m	48	0%	94%	0%	6%	W. Melle
Hardangerfjord	59°58 N 05°41 E	8/30/2016	106-0 m	46	0%	48%	0%	52%	A. Mailli
Labrador 2	59°47 N 52°16 W	5/28/2013	1000-0 m	47	0%	91%	9%	0%	I. Smolina
Oslofjord	59°12 N 10°38 E	4/16/2016	350-0 m	51	0%	63%	31%	6%	T. Falkenhaug
Lenefjord	58°5 N 07°9 E	7/18/2016	220-0 m	45	0%	56%	11%	33%	K. Eiane
Risørfjord	58°44 N 09°15 E	4/15/2016	160-0 m	47	0%	87%	0%	13%	T. Falkenhaug
Grønsfjord	58°2 N 07°2 E	7/18/2016	190-0 m	48	0%	75%	8%	17%	K. Eiane
Loch Etive	56°27 N 05°11 W	9/19/2016	130-5 m	47	0%	100%	0%	0%	K. Last
Off-Ainort	56°51 N 05°53 W	7/31/2008	20-0 m	48	0%	4%	0%	96%	S. Wells
Quebec 1	50°26 N 50°42 W	5/11/2014	200-0 m	47	0%	77%	23%	0%	A. Bucklin
Quebec 2	45°05 N 53°44 W	5/15/2014	200-0 m	47	4%	70%	26%	0%	A. Bucklin
US Portland	42°59 N 68°52 W	8/22/2012	200-0 m	47	0%	100%	0%	0%	A. Bucklin
US New-York	40°19 N 69°50 W	8/14/2012	200-0 m	45	0%	100%	0%	0%	A. Bucklin
US Philadelphia	39°58 N 72°46 W	8/8/2012	200-0 m	51	0%	100%	0%	0%	A. Bucklin

SUPPLEMENTARY 2



Phylogenetic relationships among 16S rDNA individual sequences of *Calanus finmarchicus*, *C. glacialis*, *C. hyperboreus* and *C. helgolandicus*

SUPPLEMENTARY 3

Test of congruence of *Calanus* species identification between 6 nuclear InDel markers and mitochondrial 16S rDNA, for individuals from the North Atlantic and Arctic Oceans. 4 species identified are *C. finmarchicus* (*Cfin*), *C. glacialis* (*Cgla*), *C. hyperboreus* (*Chyp*) and *C. helgolandicus* (*Chelg*).

Sampling location	Individual name	InDels Sp ID	16S Sp ID	GenBank Accession Number sequences 16S rDNA	Congruence of markers
Balsfjord 2	Bal117	*Cfin*	*Cfin*	MF959726	yes
Balsfjord 2	Bal130	*Cfin*	*Cfin*	MF959726	yes
Bjørnøya 1	Bj051	*Cgla*	*Cgla*	MF959707	yes
Bjørnøya 1	Bj054	*Cfin*	*Cfin*	MF959726	yes
Bjørnøya 1	Bj065	*Cgla*	*Cgla*	MF959707	yes
Bjørnøya 1	Bj078	*Cfin*	*Cfin*	MF959726	yes
Bjørnøya 2	409-34	*Cfin*	*Cfin*	MF959726	yes
Bjørnøya 2	409-47	*Cfin*	*Cfin*	MF959709	yes
Christianssund	187-26	*Cfin*	*Cfin*	MF959726	yes
Chuckchi Sea	Chuck1	*Chyp*	*Chyp*	MF959727	yes
Chuckchi Sea	Chuck2	*Chyp*	*Chyp*	MF972920	yes
Grønsfjord	Gro15	*Chyp*	*Chyp*	MF972920	yes
Grønsfjord	Gro21	*Cfin*	*Cfin*	MF959726	yes
Grønsfjord	Gro23	*Cfin*	*Cfin*	MF959726	yes
Grønsfjord	Gro24	*Cfin*	*Cfin*	MF959702	yes
Grønsfjord	Gro27	*Chel*	*Chel*	MF959717	yes
Grønsfjord	Gro29	*Chel*	*Chel*	MF959717	yes
Grønsfjord	Gro3	*Chyp*	*Chyp*	MF959730	yes
Grønsfjord	Gro30	*Chyp*	*Chyp*	MF959730	yes
Grønsfjord	Gro31	*Chel*	*Chel*	MF959717	yes
Grønsfjord	Gro46	*Chyp*	*Chyp*	MF959718	yes
Isfjord	Is100	*Cgla*	*Cgla*	MF959707	yes
Isfjord	Is144	*Cgla*	*Cgla*	MF959707	yes
Isfjord	Is176	*Cgla*	*Cgla*	MF959707	yes
Isfjord	Is216	*Cgla*	*Cgla*	MF959707	yes
Isfjord	Is220	*Cfin*	*Cfin*	MF959712	yes
Korsfjord	Bergen103	*Cfin*	*Cfin*	MF959726	yes
Korsfjord	Bergen104	*Chel*	*Chel*	MF959717	yes
Korsfjord	Bergen111	*Cfin*	*Cfin*	MF959726	yes
Korsfjord	Bergen123	*Cfin*	*Cfin*	MF959726	yes
Korsfjord	Bergen144	*Cfin*	*Cfin*	MF959726	yes
Labrador 1	176-80	*Cfin*	*Cfin*	MF959726	yes
Labrador 1	176-84	*Chyp*	*Chyp*	MF959725	yes
Labrador 1	176-91	*Cgla*	*Cgla*	MF959707	yes
Laptev Sea	Lapt1	*Chyp*	*Chyp*	MF959723	yes
Laptev Sea	Lapt17	*Chyp*	*Chyp*	MF959705	yes
Laptev Sea	Lapt18	*Cgla*	*Cgla*	MF959707	yes
Laptev Sea	Lapt27	*Cfin*	*Cfin*	MF959726	yes
Lenefjord	Lene1	*Chyp*	*Chyp*	MF959730	yes
Lenefjord	Lene11	*Cfin*	*Cfin*	MF959729	yes
Lenefjord	Lene21	*Chel*	*Chel*	MF959719	yes
Lenefjord	Lene22	*Chyp*	*Chyp*	MF959713	yes
Lenefjord	Lene23	*Chyp*	*Chyp*	MF959715	yes
Lenefjord	Lene43	*Chel*	*Chel*	MF959717	yes
Lenefjord	Lene45	*Chyp*	*Chyp*	MF959730	yes

Sampling location	Individual name	InDels Sp ID	16S Sp ID	GenBank Accession Number sequences 16S rDNA	Congruence of markers
Lurefjord 1	Lure25	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 1	Lure29	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 1	Lure33	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 1	Lure50	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 2	Lure86	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 3	Lure142	*Chel*	*Chel*	MF972922	yes
Lurefjord 3	Lure144	*Cgla*	*Cgla*	MF959707	yes
Lurefjord 3	Lure149	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist11	*Cfin*	*Cfin*	MF959726	yes
Mistfjord 2	Mist12	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist16	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist18	*Cgla*	*Cgla*	MF959728	yes
Mistfjord 2	Mist19	*Chyp*	*Chyp*	MF972920	yes
Mistfjord 2	Mist21	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist22	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist31	*Chyp*	*Chyp*	MF959730	yes
Mistfjord 2	Mist36	*Chyp*	*Chyp*	MF959708	yes
Mistfjord 2	Mist39	*Chyp*	*Chyp*	MF959730	yes
Mistfjord 2	Mist44	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist45	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist47	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist48	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist5	*Cgla*	*Cgla*	MF959707	yes
Mistfjord 2	Mist50	*Cfin*	*Cfin*	MF959726	yes
Mistfjord 2	Mist52	*Chyp*	*Chyp*	MF959730	yes
Mistfjord 2	Mist88	*Cgla*	*Cgla*	MF959707	yes
Myken 1	Myken22	*Cgla*	*Cgla*	MF959707	yes
Myken 1	Myken3	*Cfin*	*Cfin*	MF959721	yes
Myken 1	Myken30	*Cfin*	*Cfin*	MF959726	yes
Myken 1	Myken47	*Chyp*	*Chyp*	MF972920	yes
Myken 2	Myken80	*Cfin*	*Cfin*	MF959726	yes
Myken 3	Myken99	*Cfin*	*Cfin*	MF959711	yes
Myken 4	Myken114	*Cfin*	*Cfin*	MF959726	yes
Myken 6	Myken130	*Chel*	*Chel*	MF959717	yes
Myken 6	Myken137	*Chel*	*Chel*	MF959717	yes
Myken 6	Myken138	*Chel*	*Chel*	MF959717	yes
North Atlantic 1	160-17	*Cgla*	*Cgla*	MF959707	yes
North Atlantic 1	160-2	*Chyp*	*Chyp*	MF972920	yes
North Atlantic 2	155-29	*Cgla*	*Cgla*	MF959707	yes
North Atlantic 3	153-15	*Cfin*	*Cfin*	MF959726	yes
North Atlantic 3	153-45	*Cfin*	*Cfin*	MF959726	yes
North Atlantic 4	168-20	*Cfin*	*Cfin*	MF959726	yes
North Iceland	165-29	*Cfin*	*Cfin*	MF959726	yes
North Sea	491-39	*Chel*	*Chel*	MF959717	yes
Off-Tromsø	Tromso004	*Cfin*	*Cfin*	MF959726	yes
Off-Tromsø	Tromso035	*Cfin*	*Cfin*	MF959726	yes
Off-Tromsø	Tromso20	*Chel*	*Chel*	MF959717	yes
Off-Tromsø	Tromso91	*Chel*	*Chel*	MF959724	yes
Oslofjord	Oslo14	*Chyp*	*Chyp*	MF959730	yes
Oslofjord	Oslo15	*Chyp*	*Chyp*	MF959730	yes
Oslofjord	Oslo16	*Chyp*	*Chyp*	MF959716	yes
Oslofjord	Oslo17	*Chyp*	*Chyp*	MF959730	yes
Oslofjord	Oslo24	*Cfin*	*Cfin*	MF959726	yes
Oslofjord	Oslo27	*Chel*	*Chel*	MF959717	yes
Oslofjord	Oslo42	*Chel*	*Chel*	MF959717	yes
Oslofjord	Oslo49	*Cfin*	*Cfin*	MF959726	yes

Sampling location	Individual name	InDels Sp ID	16S Sp ID	GenBank Accession Number sequences 16S rDNA	Congruence of markers
Østerbotn	Por68	*Cfin*	*Cfin*	MF959726	yes
Østerbotn	Por83	*Cgla*	*Cgla*	MF959707	yes
Osterfjord	Ost17	*Chel*	*Chel*	MF972922	yes
Osterfjord	Ost23	*Cfin*	*Cfin*	MF959726	yes
Porsangerfjord 1	Por46	*Cfin*	*Cfin*	MF959726	yes
Ranfjord 1	Ran6	*Cgla*	*Cgla*	MF959707	yes
Ranfjord 1	Ran7	*Cgla*	*Cgla*	MF959706	yes
Ranfjord 3	Ran119	*Cfin*	*Cfin*	MF959726	yes
Ranfjord 3	Ran124	*Cfin*	*Cfin*	MF959726	yes
Ranfjord 3	Ran128	*Cfin*	*Cfin*	MF959726	yes
Ranfjord 3	Ran60	*Cfin*	*Cfin*	MF959726	yes
Raunefjord	Bergen19	*Cfin*	*Cfin*	MF959726	yes
Raunefjord	Bergen7	*Cfin*	*Cfin*	MF959726	yes
Risørfjord	Ris15	*Chel*	*Chel*	MF959720	yes
Risørfjord	Ris13	*Chel*	*Chel*	MF959717	yes
Saltfjord 1	150216-095	*Cgla*	*Cgla*	MF959707	yes
Saltfjord 1	150216-105	*Cgla*	*Cgla*	MF959707	yes
Saltfjord 1	150216-31	*Chyp*	*Chyp*	MF959730	yes
Saltfjord 2	Salt25	*Cgla*	*Cgla*	MF959707	yes
Saltfjord 2	Salt53	*Cfin*	*Cfin*	MF959726	yes
Skjerstadfjord 1	Skj006	*Chyp*	*Chyp*	MF959730	yes
Skjerstadfjord 1	Skj24	*Chyp*	*Chyp*	MF959710	yes
Skjerstadfjord 1	Skj41	*Cfin*	*Cfin*	MF959729	yes
Skjerstadfjord 1	Skj44	*Cgla*	*Cgla*	MF959707	yes
Skjerstadfjord 1	Skj46	*Cgla*	*Cgla*	MF959707	yes
Sørfjord 1	Sorfjord14	*Cfin*	*Cfin*	MF959726	yes
Sørfjord 2	Sorfjord105	*Cfin*	*Cfin*	MF959726	yes
Sørfjord 2	Sorfjord112	*Chel*	*Chel*	MF959717	yes
Sørfjord 2	Sorfjord27	*Chel*	*Chel*	MF972922	yes
Sørfjord 2	Sorfjord42	*Chel*	*Chel*	MF959717	yes
Sørfjord 3	Sorfjord128	*Cfin*	*Cfin*	MF959726	yes
Sørfoldfjord 1	Sorfolda61	*Cgla*	*Cgla*	MF959707	yes
Sørfoldfjord 1	Sorfolda63	*Cgla*	*Cgla*	MF959707	yes
Sørfoldfjord 1	Sorfolda64	*Cgla*	*Cgla*	MF959707	yes
Sørfoldfjord 1	Sorfolda87	*Cfin*	*Cfin*	MF959702	yes
Sørfoldfjord 2	Sorfolda108	*Chyp*	*Chyp*	MF972920	yes
Sørfoldfjord 2	Sorfolda143	*Chyp*	*Chyp*	MF959730	yes
Sørfoldfjord 3	Sorfolda158	*Chyp*	*Chyp*	MF959730	yes
Sørfoldfjord 3	Sorfolda187	*Cgla*	*Cgla*	MF959707	yes
Steinkjer	Stein18	*Chel*	*Chel*	MF959717	yes
Steinkjer	Stein19	*Cfin*	*Cfin*	MF959729	yes
Steinkjer	Stein20	*Cfin*	*Cfin*	MF959729	yes
Svinøy	Sv017	*Chel*	*Chel*	MF972921	yes
Svinøy	Sv077	*Chel*	*Chel*	MF959720	yes
Svinøy	Sv70	*Chel*	*Chel*	MF959717	yes
Svinøy	Sv83	*Chel*	*Chel*	MF959722	yes
Tronsdheimsfjord	Tdh26	*Cfin*	*Cfin*	MF959726	yes
Tysfjord	Tys20	*Chyp*	*Chyp*	MF959730	yes
Tysfjord	Tys4	*Chyp*	*Chyp*	MF959710	yes
Van Mijenfjord	VM66	*Chyp*	*Chyp*	MF959714	yes
Vestfjord 1	Lof29	*Cgla*	*Cgla*	MF959707	yes
Vestfjord 2	Vest10	*Cgla*	*Cgla*	MF959707	yes
Vestfjord 2	Vest22	*Cgla*	*Cgla*	MF959707	yes
Vestfjord 2	Vest29	*Cgla*	*Cgla*	MF959707	yes
Vestfjord 3	Vest63	*Cgla*	*Cgla*	MF959707	yes
Vestfjord 3	Vest92	*Cgla*	*Cgla*	MF959707	yes
White Sea	WS2	*Cgla*	*Cgla*	MF959707	yes
White Sea	WS34	*Cgla*	*Cgla*	MF959707	yes

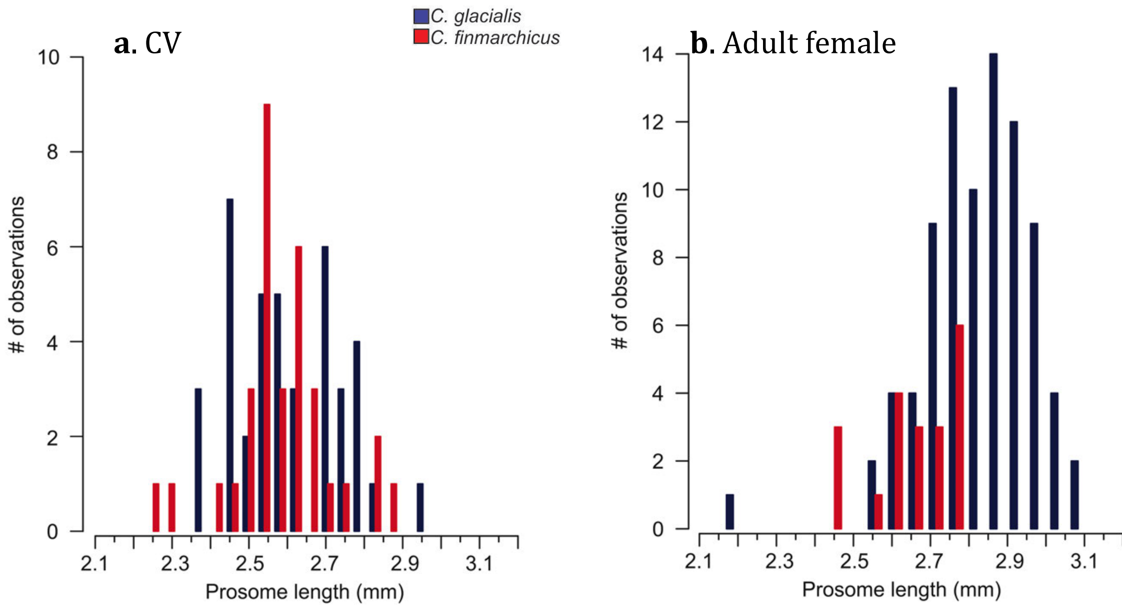
SUPPLEMENTARY 4

Test of congruence of *Calanus* species identification between 6 nuclear InDel markers and mitochondrial 16S rDNA for *Calanus finmarchicus* (*Cfin*) and *C. glacialis* (*Cgla*) in Saltenfjord/Skjerstadfjord accompanied with prosome length measurements and developmental stage information.

Sampling location	Individual name	InDels Sp ID	16S Sp ID	GenBank Accession Number sequences 16S rDNA	Congruence of markers	Prosome Length (µm)	Developmental Stage
Saltfjord 1	15-02-16-009	*Cgla*	*Cgla*	MF959707	yes	2303,17	CV
Saltfjord 1	15-02-16-011	*Cgla*	*Cgla*	MF959707	yes	2362,44	CV
Saltfjord 1	15-02-16-012	*Cgla*	*Cgla*	MF959706	yes	2132,79	CV
Saltfjord 1	15-02-16-013	*Cfin*	*Cfin*	MF959729	yes	2403,55	CV
Saltfjord 1	15-02-16-015	*Cgla*	*Cgla*	MF959706	yes	2383,01	CV
Saltfjord 1	15-02-16-016	*Cfin*	*Cfin*	MF959726	yes	2379,29	CV
Saltfjord 1	15-02-16-018	*Cgla*	*Cgla*	MF959706	yes	2502,75	CV
Saltfjord 1	15-02-16-019	*Cfin*	*Cfin*	MF959726	yes	2413,01	CV
Saltfjord 1	15-02-16-020	*Cfin*	*Cfin*	MF959726	yes	2482,57	CV
Saltfjord 1	15-02-16-021	*Cfin*	*Cfin*	MF959703	yes	2644,03	Female
Saltfjord 1	15-02-16-024	*Cfin*	*Cfin*	MF959729	yes	2316,18	CV
Saltfjord 1	15-02-16-025	*Cfin*	*Cfin*	MF959726	yes	2282,84	CV
Saltfjord 1	15-02-16-026	*Cgla*	*Cgla*	MF959707	yes	2405,64	CV
Saltfjord 1	15-02-16-028	*Cfin*	*Cfin*	MF959726	yes	2717,74	Female
Saltfjord 1	15-02-16-030	*Cgla*	*Cgla*	MF959707	yes	2609,69	CV
Saltfjord 1	15-02-16-032	*Cgla*	*Cgla*	MF959707	yes	2607,77	CV
Saltfjord 1	15-02-16-033	*Cgla*	*Cgla*	MF959707	yes	2226,31	CV
Saltfjord 1	15-02-16-034	*Cfin*	*Cfin*	MF959726	yes	2426,94	CV
Saltfjord 1	15-02-16-035	*Cfin*	*Cfin*	MF959726	yes	2669,51	CV
Saltfjord 1	15-02-16-036	*Cgla*	*Cgla*	MF959707	yes	2309,67	CV
Saltfjord 1	15-02-16-037	*Cfin*	*Cfin*	MF959726	yes	2236,15	CV
Saltfjord 1	15-02-16-039	*Cfin*	*Cfin*	MF959726	yes	2521,78	Female
Saltfjord 1	15-02-16-040	*Cfin*	*Cfin*	MF959726	yes	2328,63	CV
Saltfjord 1	15-02-16-042	*Cfin*	*Cfin*	MF959726	yes	2265,14	CV
Saltfjord 1	15-02-16-043	*Cfin*	*Cfin*	MF959726	yes	2186,02	CV
Saltfjord 1	15-02-16-044	*Cfin*	*Cfin*	MF959726	yes	2339,14	CV
Saltfjord 1	15-02-16-045	*Cfin*	*Cfin*	MF959726	yes	2266,17	CV
Saltfjord 1	15-02-16-046	*Cfin*	*Cfin*	MF959726	yes	2333,27	CV
Saltfjord 1	15-02-16-047	*Cfin*	*Cfin*	MF959726	yes	2301,69	CV
Saltfjord 1	15-02-16-048	*Cgla*	*Cgla*	MF959707	yes	2224,37	CV
Saltfjord 1	15-02-16-049	*Cgla*	*Cgla*	MF959706	yes	2359,86	CV
Saltfjord 1	15-02-16-050	*Cgla*	*Cgla*	MF959706	yes	2119,4	CV
Saltfjord 1	15-02-16-051	*Cfin*	*Cfin*	MF959726	yes	2691,54	CV
Saltfjord 1	15-02-16-052	*Cgla*	*Cgla*	MF959706	yes	2348,86	CV
Saltfjord 1	15-02-16-053	*Cfin*	*Cfin*	MF959726	yes	2320,71	CV
Saltfjord 1	15-02-16-054	*Cfin*	*Cfin*	MF959726	yes	2747,02	Female
Saltfjord 1	15-02-16-056	*Cfin*	*Cfin*	MF959726	yes	2555,2	Female
Saltfjord 1	15-02-16-058	*Cgla*	*Cgla*	MF959707	yes	2583,62	CV
Saltfjord 1	15-02-16-059	*Cfin*	*Cfin*	MF959726	yes	2367,53	CV
Saltfjord 1	15-02-16-061	*Cfin*	*Cfin*	MF959726	yes	2416,58	Female
Saltfjord 1	15-02-16-062	*Cfin*	*Cfin*	MF959726	yes	2725,01	Female
Saltfjord 1	15-02-16-064	*Cfin*	*Cfin*	MF959726	yes	2717,76	CV
Saltfjord 1	15-02-16-066	*Cfin*	*Cfin*	MF959726	yes	2031,03	CV
Saltfjord 1	15-02-16-067	*Cgla*	*Cgla*	MF959707	yes	2519,75	CV
Saltfjord 1	15-02-16-068	*Cgla*	*Cgla*	MF959706	yes	2219,32	CV
Skjerstadfjord 2	26-02-16-002	*Cgla*	*Cgla*	MF959707	yes	2819,62	Female
Skjerstadfjord 2	26-02-16-003	*Cgla*	*Cgla*	MF959707	yes	2866,97	Female
Skjerstadfjord 2	26-02-16-004	*Cfin*	*Cfin*	MF959726	yes	2487,92	CV
Skjerstadfjord 2	26-02-16-008	*Cgla*	*Cgla*	MF959707	yes	2662,64	Female
Skjerstadfjord 2	26-02-16-010	*Cgla*	*Cgla*	MF959707	yes	2745,27	Female
Skjerstadfjord 2	26-02-16-011	*Cgla*	*Cgla*	MF959707	yes	2682,25	Female
Skjerstadfjord 2	26-02-16-012	*Cgla*	*Cgla*	MF959707	yes	2770,09	Female
Skjerstadfjord 2	26-02-16-014	*Cgla*	*Cgla*	MF959707	yes	2772,95	Female
Skjerstadfjord 2	26-02-16-015	*Cgla*	*Cgla*	MF959707	yes	2738,5	Female
Skjerstadfjord 2	26-02-16-016	*Cgla*	*Cgla*	MF959707	yes	2150,68	Female
Skjerstadfjord 2	26-02-16-017	*Cgla*	*Cgla*	MF959707	yes	2958,1	Female
Skjerstadfjord 2	26-02-16-018	*Cgla*	*Cgla*	MF959707	yes	2819,35	Female

Sampling location	Individual name	InDels Sp ID	16S Sp ID	GenBank Accession Number sequences 16S rDNA	Congruence of markers	Prosome Length (μm)	Developmental Stage
Skjerstadfjord 2	26-02-16-019	*Cgla*	*Cgla*	MF959707	yes	2550,68	CV
Skjerstadfjord 2	26-02-16-022	*Cgla*	*Cgla*	MF959707	yes	2868,47	Female
Skjerstadfjord 2	26-02-16-023	*Cgla*	*Cgla*	MF959707	yes	2968,3	Female
Skjerstadfjord 2	26-02-16-026	*Cgla*	*Cgla*	MF959707	yes	2767,48	Female
Skjerstadfjord 2	26-02-16-027	*Cgla*	*Cgla*	MF959707	yes	2736,55	Female
Skjerstadfjord 2	26-02-16-028	*Cgla*	*Cgla*	MF959707	yes	2937,8	Female
Skjerstadfjord 2	26-02-16-030	*Cgla*	*Cgla*	MF959707	yes	2683,26	Female
Skjerstadfjord 2	26-02-16-031	*Cgla*	*Cgla*	MF959707	yes	2812,78	Female
Skjerstadfjord 2	26-02-16-032	*Cgla*	*Cgla*	MF959707	yes	2854,29	Female
Skjerstadfjord 2	26-02-16-033	*Cgla*	*Cgla*	MF959707	yes	2748,12	Female
Skjerstadfjord 2	26-02-16-036	*Cgla*	*Cgla*	MF959707	yes	2790,18	Female
Skjerstadfjord 2	26-02-16-037	*Cgla*	*Cgla*	MF959707	yes	2805,97	Female
Skjerstadfjord 2	26-02-16-039	*Cgla*	*Cgla*	MF959707	yes	2727,74	Female
Skjerstadfjord 2	26-02-16-044	*Cgla*	*Cgla*	MF959707	yes	2722,62	Female
Skjerstadfjord 2	26-02-16-045	*Cgla*	*Cgla*	MF959707	yes	2219,54	CV
Skjerstadfjord 2	26-02-16-046	*Cgla*	*Cgla*	MF959707	yes	2529,42	Female
Skjerstadfjord 2	26-02-16-047	*Cgla*	*Cgla*	MF959707	yes	2754,26	Female
Skjerstadfjord 2	26-02-16-048	*Cgla*	*Cgla*	MF959707	yes	2931,09	Female
Skjerstadfjord 2	26-02-16-049	*Cgla*	*Cgla*	MF959707	yes	2712,63	Female
Skjerstadfjord 2	26-02-16-050	*Cgla*	*Cgla*	MF959707	yes	2861,56	Female
Skjerstadfjord 2	26-02-16-051	*Cgla*	*Cgla*	MF959707	yes	2709,82	Female
Skjerstadfjord 2	26-02-16-052	*Cgla*	*Cgla*	MF959707	yes	2900,9	Female
Skjerstadfjord 2	26-02-16-054	*Cgla*	*Cgla*	MF959707	yes	2834,98	Female
Skjerstadfjord 2	26-02-16-055	*Cgla*	*Cgla*	MF959707	yes	2848,86	Female
Skjerstadfjord 2	26-02-16-056	*Cgla*	*Cgla*	MF959707	yes	2935,77	Female
Skjerstadfjord 2	26-02-16-057	*Cgla*	*Cgla*	MF959707	yes	2775,33	Female
Skjerstadfjord 2	26-02-16-060	*Cgla*	*Cgla*	MF959707	yes	2830	Female
Skjerstadfjord 2	26-02-16-065	*Cgla*	*Cgla*	MF959707	yes	2707,59	Female
Skjerstadfjord 2	26-02-16-066	*Cgla*	*Cgla*	MF959707	yes	2777,23	Female
Skjerstadfjord 2	26-02-16-067	*Cfin*	*Cfin*	MF959704	yes	2738,81	Female
Skjerstadfjord 2	26-02-16-069	*Cgla*	*Cgla*	MF959707	yes	2935,35	Female
Skjerstadfjord 2	26-02-16-070	*Cgla*	*Cgla*	MF959707	yes	2884,35	Female
Skjerstadfjord 2	26-02-16-071	*Cgla*	*Cgla*	MF959707	yes	2852,28	Female
Skjerstadfjord 2	26-02-16-072	*Cgla*	*Cgla*	MF959707	yes	2560,02	Female
Skjerstadfjord 2	26-02-16-074	*Cfin*	*Cfin*	MF959726	yes	2623,18	Female
Skjerstadfjord 2	26-02-16-075	*Cfin*	*Cfin*	MF959726	yes	2440,78	CV
Skjerstadfjord 2	26-02-16-076	*Cfin*	*Cfin*	MF959726	yes	2686,92	Female
Skjerstadfjord 2	26-02-16-078	*Cgla*	*Cgla*	MF959707	yes	2909,09	Female
Skjerstadfjord 2	26-02-16-079	*Cfin*	*Cfin*	MF959726	yes	2446,4	CV
Skjerstadfjord 2	26-02-16-080	*Cgla*	*Cgla*	MF959707	yes	2845,11	Female
Skjerstadfjord 2	26-02-16-081	*Cgla*	*Cgla*	MF959707	yes	2771,56	Female
Skjerstadfjord 2	26-02-16-083	*Cgla*	*Cgla*	MF959707	yes	2916,63	Female
Skjerstadfjord 2	26-02-16-085	*Cfin*	*Cfin*	MF959726	yes	2548,65	CV
Skjerstadfjord 2	26-02-16-087	*Cgla*	*Cgla*	MF959707	yes	2946,72	Female
Skjerstadfjord 2	26-02-16-089	*Cgla*	*Cgla*	MF959706	yes	2817,67	Female
Skjerstadfjord 2	26-02-16-090	*Cgla*	*Cgla*	MF959706	yes	2893,8	Female
Skjerstadfjord 2	26-02-16-091	*Cgla*	*Cgla*	MF959707	yes	2850,28	Female
Skjerstadfjord 2	26-02-16-092	*Cgla*	*Cgla*	MF959707	yes	2871,76	Female
Skjerstadfjord 2	26-02-16-093	*Cgla*	*Cgla*	MF959707	yes	2747	Female
Skjerstadfjord 2	26-02-16-094	*Cgla*	*Cgla*	MF959707	yes	2811,26	Female
Skjerstadfjord 2	26-02-16-095	*Cgla*	*Cgla*	MF959707	yes	2700,08	Female
Skjerstadfjord 2	26-02-16-096	*Cgla*	*Cgla*	MF959707	yes	2596,9	Female
Skjerstadfjord 2	26-02-16-098	*Cgla*	*Cgla*	MF959707	yes	2859,26	Female
Skjerstadfjord 2	26-02-16-132	*Cgla*	*Cgla*	MF959707	yes	2886,28	Female
Skjerstadfjord 2	26-02-16-134	*Cgla*	*Cgla*	MF959707	yes	2837,12	Female
Skjerstadfjord 2	26-02-16-135	*Cfin*	*Cfin*	MF959726	yes	2694,29	Female
Skjerstadfjord 2	26-02-16-138	*Cgla*	*Cgla*	MF959707	yes	2750,05	Female
Skjerstadfjord 2	26-02-16-139	*Cfin*	*Cfin*	MF959702	yes	2416,67	Female
Skjerstadfjord 2	26-02-16-140	*Cgla*	*Cgla*	MF959707	yes	2623,33	CV
Skjerstadfjord 2	26-02-16-141	*Cfin*	*Cfin*	MF959726	yes	1976,64	CV
Skjerstadfjord 2	26-02-16-142	*Cgla*	*Cgla*	MF959707	yes	3030,50	Female
Skjerstadfjord 2	26-02-16-144	*Cgla*	*Cgla*	MF959707	yes	2652,26	Female
Skjerstadfjord 2	26-02-16-145	*Cgla*	*Cgla*	MF959707	yes	2964,65	Female
Skjerstadfjord 2	26-02-16-146	*Cgla*	*Cgla*	MF959707	yes	2521,83	Female
Skjerstadfjord 2	26-02-16-147	*Cgla*	*Cgla*	MF959707	yes	2669,51	Female
Skjerstadfjord 2	26-02-16-149	*Cfin*	*Cfin*	MF959726	yes	2620,01	Female
Skjerstadfjord 2	26-02-16-150	*Cfin*	*Cfin*	MF959726	yes	2406,89	Female
Skjerstadfjord 2	26-02-16-152	*Cgla*	*Cgla*	MF959707	yes	2823	Female
Skjerstadfjord 2	26-02-16-153	*Cgla*	*Cgla*	MF959707	yes	2782,33	Female
Skjerstadfjord 2	26-02-16-154	*Cgla*	*Cgla*	MF959707	yes	2621,5	Female
Skjerstadfjord 2	26-02-16-156	*Cgla*	*Cgla*	MF959707	yes	2965,81	Female
Skjerstadfjord 2	26-02-16-157	*Cgla*	*Cgla*	MF959707	yes	2933,8	Female

SUPPLEMENTARY 5



Frequency distributions of prosome length for *Calanus glacialis* and *C. finmarchicus* at developmental stage (a) CV and (b) adult female in the region of Saltenfjord / Skjerstadfjord. In total, prosome length of 171 *Calanus* individuals was measured.

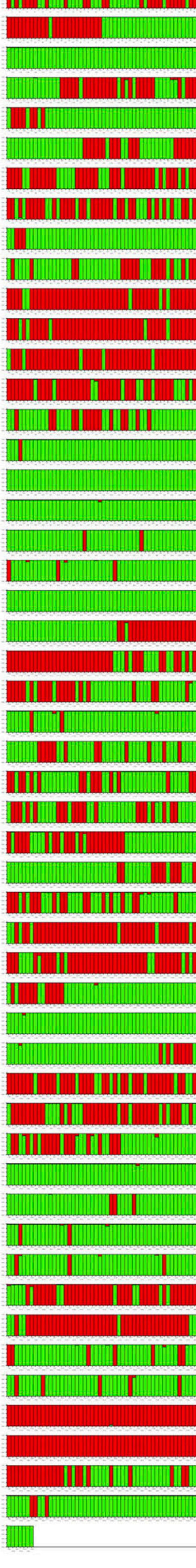
SUPPLEMENTARY 6

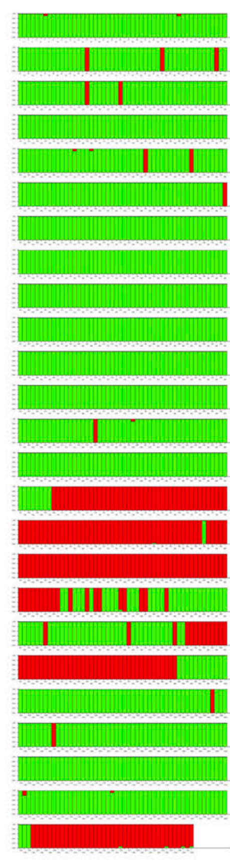
Genotype admixture analysis based on nuclear InDels shows no hybrids between *Calanus finmarchicus* and *C. glacialis*

Bar chart representing genotype admixture analysis results based on nuclear InDel genotypes data, performed using STRUCTURE (v. 2.3.4) (Pritchard et al., 2000), after Nielsen *et al.* (2014).

a. Results from the first set of InDels used (c.f. Supplementary7),
b. Results from the second set of InDels.

Each individual is represented by a bar filled with one or two distinct colours that identify an individual probability to belong to two clusters (here, green for *C. finmarchicus* and red for *C. glacialis*). In case of F1 hybrids between *C. finmarchicus* and *C. glacialis*, a bar will be nearly equally filled with both colours (which never happened).





SUPPLEMENTARY 7

Supplementary Protocols

- DNA extraction:

We extracted DNA from the antennules of each specimen, using the quick and cheap method of HotSHOT DNA extraction (Montero-Pau et al., 2008):

1- Individuals were soaked separately in sterile water to rinse the ethanol;
2- One by one, under a stereomicroscope, the 2 antennules were removed from the rest of the body and placed in 50 µL of a Lysis Buffer (See HotSHOT protocol for details about composition of buffers and Montero-Pau et al., 2008) in a 96-well plate;
3- The plate was incubed in a thermocycler, 30 minutes at 95°C;
4- The plate was subsequently cooled in the fridge (4°C) for 5-10 minutes;
5- Finally, 50 µL of Neutralizing Solution was added (See HotSHOT protocol for details about composition of buffers (Montero-Pau et al., 2008)).

- Molecular species identification:

We amplified a set of 6 nuclear molecular markers, type InDel (polymorphism consists of Insertion or Deletion of nucleotides): G_150, T_461, T_1338, T_1966, T_3133 and T_4700 (Smolina et al., 2014) in a single multiplexed Polymerase Chain Reaction (PCR), and genotyped them following the protocol described by Smolina *et al*. (2014). Four distinct patterns of genotypes were distinguished and assigned to the four different species of *Calanus* based on species-specific alleles defined in Smolina *et al*. (2014). This method is fast and inexpensive. A total of 96 individuals can be reliably identified within 5 hours, with 100% reliable results for ca. 2 euros/sample. At one point in our study, we had to re-order a new stock of InDel primers from a new provider, and thus had to change the type of fluorescent dye labelling of the forward primers (from 6-FAM, VIC and NED (Life Technologies) to FAM, YAKYE and ATTO550 (Eurofins Genomics)). This resulted in a slight shift of the length of the alleles in the genotyping, thus this second set of data was treated separately. To confirm the species identification, and in order to validate our nuclear markers, we sequenced a portion of the mitochondrial 16S rDNA (ca. 360bp) (Sanger and Coulson, 1975) for 159 individuals from 53 locations selected to represent the full range of sampling, and for 129 individuals from the region of Saltenfjord / Skjerstadfjord, following the same protocol described in Smolina *et al.* (2014). The obtained sequences were then aligned together with one reference 16S sequence for each species from GenBank®: HQ266740 for *C. glacialis*, AF295334 for *C. finmarchicus*, KF956849 for *C. helgolandicus*, and JX678968 for *C. hyperboreus*. This alignment was used to reconstruct a PhyML tree (GTR model) using Geneious version 9.1 (http://www.geneious.com) (Kearse et al., 2012). The resulting tree displayed four clearly distinct groups of sequences corresponding to the four species (see Supplementary2). In all individuals, this approach resulted in the same species identification as the InDel genotyping (Supplementary3-4).

- Microsatellite analysis:

To characterize connectivity among newly described population of *C. glacialis* in Norwegian fjords and other regions and compare it to *C. finmarchicus* we performed analysis of population genetic differentiation using sample from 3 locations: Isfjord, Saltfjord and Lurefjord (c.f. Supplementary1). DNA from the antennas of 24 identified (InDels method – see above) individuals per species and per location was used to amplify 10 microsatellites markers (Parent et al., 2012, Provan et al., 2009) by PCR. Nine microsatellite markers were multiplexed into 3 PCR reactions (EL696609, EL585922, and EL773519; FK868270, FG632811, and FK670364; EH666870, EH666474, and EL773359) and one marker, FK867682, was amplified separately. PCR reactions were carried out in a final volume of 5 µL, using 2.5 µL of Accu-Start Tough Mix® (Quanta Biosciences), 0.1 µL of each primer at 10 µM, and completing with DNA from HotSHOT extraction. Reactions were run on a Veriti® 96-well fast thermal cycler (Applied Biosystems). Amplification reactions consisted of Multiplex 1 and FK867682 for 3 min at 95°C; with 35 cycles of: 95°C (30 sec), 55°C (60 sec), and 72°C (60 sec); with one final extension cycle at 72°C for 5 min. Multiplex 2 was amplified by a first cycle at 95°C (3 min), followed by 35 cycles of: 95°C (30 sec), 53°C (60 sec), and 72°C (60 sec); and one final cycle at 72°C (5 min) for extension. Multiplex 3 was run as 95°C (3 min), and then 10 cycles of: 95°C (1 min), touch-down from 68°C (1 min) to 58°C (1 min), and 72°C (1 min); followed by 25 cycles of: 95°C (1 min), 58°C (1 min), and 72°C (1 min); and a final extension cycle at 72°C (5 min). Microsatellite loci were analysed on a 3500XL Genetic Analyzer (Life Technologies) in three multiplex and one singleplex following Nielsen *et al.* (2014).

We used *Genetix* (v. 4.05.2) (Belkhir et al., 1996) to estimate the global F_{ST}, index of population differentiation after Weir & Cockerham (1984).

SUPPLEMENTARY 8

Sources of literature used for tracing the morphologically based distribution ranges of *Calanus* species

The map showing the distribution ranges of *Calanus* species in the North Atlantic and Arctic Ocean, as defined from morphological identification of species, presented as Figure 1 of the paper, was mainly based on three different sources: Conover, 1988, Barnard et al., 2004 for the southern borders of species distributions, and Jaschnov, 1970 for the northern borders. However, the map was completed in regard to other sources of the existing literature, and therefore the following were used as a complementary support: Wassmann et al., 2015; Melle et al., 2014; Estrada et al., 2012; Bonnet et al., 2005.

References:

Barnard, R., Batten, S., Beaugrand, G., Buckland, C., Conway, D., Edwards, M., Finlayson, J., Gregory, L., Halliday, N. & John, A. (2004). Continuous plankton records: Plankton atlas of the North Atlantic Ocean (1958-1999). II. Biogeographical charts. *Marine Ecology-Progress Series*: 11-75.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (1996). GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire génome, populations, interactions, CNRS UMR,* 5000: 1996-2004.

Bonnet, D., Richardson, A., Harris, R., Hirst, A., Beaugrand, G., Edwards, M., Ceballos, S., Diekman, R., López-Urrutia, A. & Valdes, L. (2005). An overview of *Calanus helgolandicus* ecology in European waters. *Progress in Oceanography,* 65: 1-53.

Conover, R. (1988). Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. *Biology of Copepods.* Springer.

Estrada, R., Harvey, M., Gosselin, M., Starr, M., Galbraith, P.S. & Straneo, F. (2012). Late-summer zooplankton community structure, abundance, and distribution in the Hudson Bay system (Canada) and their relationships with environmental conditions, 2003–2006. *Progress in oceanography,* 101: 121-145.

Jaschnov, W. (1970). Distribution of *Calanus* species in the seas of the northern hemisphere. *Internationale Revue der gesamten Hydrobiologie und Hydrographie,* 55: 197-212.

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

Melle, W., Runge, J., Head, E., Plourde, S., Castellani, C., Licandro, P., Pierson, J., Jonasdottir, S., Johnson, C. & Broms, C. (2014). The North Atlantic Ocean as habitat for *Calanus finmarchicus*: Environmental factors and life history traits. *Progress in Oceanography,* 129: 244-284.

Montero-Pau, J., Gómez, A. & Muñoz, J. (2008). Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods,* 6: 218-222.

Nielsen, T.G., Kjellerup, S., Smolina, I., Hoarau, G. & Lindeque, P. (2014). Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phenological differences? *Marine Biology,* 161: 1299-1306. DOI 10.1007/s00227-014-2419-5.

Parent, G.J., Plourde, S. & Turgeon, J. (2012). Natural hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and Northwest Atlantic. *Limnology and Oceanography,* 57: 1057-1066. DOI 10.4319/lo.2012.57.4.1057.

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

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A. & Savidge, G. (2009). High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus*. *Proceedings Biological sciences / The Royal Society,* 276: 301-7. DOI 10.1098/rspb.2008.1062.

Sanger, F. & Coulson, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology,* 94: 441IN19447-446IN20448.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14: 1072-9. DOI 10.1111/1755-0998.12241.

Wassmann, P., Kosobokova, K., Slagstad, D., Drinkwater, K., Hopcroft, R., Moore, S., Ellingsen, I., Nelson, R., Carmack, E. & Popova, E. (2015). The contiguous domains of Arctic Ocean advection: trails of life and death. *Progress in Oceanography,* 139: 42-65.

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

Paper II

Can morphology reliably distinguish between the copepods *Calanus finmarchicus* and *C. glacialis*, or is DNA the only way?

Marvin Choquet [1*], Ksenia Kosobokova [2], Sławomir Kwaśniewski [3], Maja Hatlebakk[1,4], Anusha K. S. Dhanasiri [1], Webjørn Melle [5], Malin Daase [6], Camilla Svensen [6], Janne E. Søreide [4], Galice Hoarau [1]

Affiliations:

[1] *Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway.*

[2] *P.P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia.*

[3] *Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland.*

[4] *Department of Arctic Biology, The University Centre in Svalbard, Longyearbyen, Norway.*

[5] *Institute of Marine Research, Bergen, Norway.*

[6] *Faculty of Biosciences, Fisheries and Economics, Department of Arctic and Marine Biology, UiT the Arctic University of Norway, Tromsø, Norway.*

***Corresponding author:**

Marvin Choquet – marvin.choquet@nord.no

Faculty of Biosciences and Aquaculture, Nord University

8049 Bodø, Norway

Keywords: *Calanus* – species identification – morphology – DNA – molecular markers

Abstract

Copepods of the genus *Calanus* play a key role in marine food webs as consumers of primary producers and as prey for many commercially important marine species. Within the genus, *C. glacialis* and *C. finmarchicus* are considered indicator species for Arctic and Atlantic waters, respectively, and changes in their distributions are frequently used as a tool to track climate change effects in the marine ecosystems of the northern hemisphere. Despite the extensive literature available, discrimination between these two species remains challenging. Using genetically identified individuals, we simultaneously checked the morphological characters in use for *C. glacialis* and *C. finmarchicus* identification to compare the results of molecular and morphological identification. We studied the prosome length (1); the antennules and the genital somite pigmentation (2); the morphology of the fifth pair of swimming legs and of the mandible (3). Our results show that none of these morphological criteria can reliably distinguish between *C. glacialis* and *C. finmarchicus.* This has severe implications for our current understanding of plankton ecology as a large part of our knowledge of *Calanus* may be biased due to species misidentification and may subsequently require reinvestigation with the systematic use of molecular tools.

Introduction

Copepods of the genus *Calanus* are the dominant component of the zooplankton in the North Atlantic and the Arctic (Conover, 1988, Fleminger and Hulsemann, 1977, Jaschnov, 1972, Kosobokova, 2012, Kosobokova et al., 2011) and are by far the most studied zooplankton species, with ca. 100 scientific publications per year for the last 30 years (*Web of Science*). They play a key role in marine food webs as consumers of primary producers and microzooplankton and as prey for many commercially and non-commercially important species (Beaugrand et al., 2003, Falk-Petersen et al., 2009, Gislason and Astthorsson, 2002, Michaud and Taggart, 2007, Skjoldal, 2004, Steen et al., 2007, Varpe et al., 2005). Furthermore, they are key drivers of the vertical export of material from the upper part of the water column due to the ability of packing organic material into large fast-sinking fecal pellets (Wilson et al., 2008). In marine food webs, *Calanus* spp. are essential agents of matter and energy transfer between phyto- and microzooplankton and higher trophic levels.

In the North Atlantic and Arctic regions, the Arctic species *C. glacialis* and the smaller north Atlantic *C. finmarchicus* account for most of the zooplankton biomass (Blachowiak-Samolyk, 2008, Fleminger and Hulsemann, 1977, Hassel, 1986, Kosobokova, 2012, Kosobokova and Hirche, 2009, Søreide et al., 2008). The spatial distribution of these two copepods is linked to the distribution of Arctic and Atlantic waters, respectively, and they are thus considered indicator species for these water masses (Beaugrand, 2007, Blachowiak-Samolyk, 2008, Bonnet and Frid, 2004, Daase and Eiane, 2007, Helaouët and Broms et al., 2009, Jaschnov, 1970, Jaschnov, 1966, Unstad and Tande, 1991). Recently, *C. glacialis* and *C. finmarchicus* have been regarded as beacons of climate change (Hays et al., 2005, Wassmann et al., 2015), as changes in their distribution are interpreted as changes in Atlantic water circulation and potential “Atlantification” of the Arctic (Falk-Petersen et al., 2007, Wassmann et al., 2006).

The ecological importance of *C. finmarchicus* and *C. glacialis* is unquestionable, but distinguishing between them in regions of co-occurrence has always been challenging (Hirche et al., 1994, Unstad and Tande, 1991). Three main morphological characters have been used, 1) prosome length; 2) redness of antennules and genital somite (the two spermathecae); 3) structure of the fifth pair of swimming legs and the coxal endid of the mandible (in adults).

Because of convenience, the prosome length measurements (1) have been and remain the most commonly used method to separate the two species (see for example: Arnkværn et al., 2005, Forest et al., 2011, Hirche and Kosobokova, 2011, Kosobokova, 2012, Kwasniewski et al., 2003, Unstad and Tande, 1991) although

several recent studies have demonstrated a size-overlap in specific regions (Gabrielsen et al., 2012, Lindeque et al., 2006, Parent et al., 2011).

Another trait that has been recently suggested to distinguish between *C. finmarchicus* and *C. glacialis* is the presence or absence of red pigmentation on their antennules and, in the case of adult females, on their genital somite (originally genital field) (2) (Nielsen et al., 2014). Examination of this character requires that individuals are alive, so the samples have to be sorted directly after collection, which is also a challenge.

The classical, but most complex and time-consuming approach to identify *C. finmarchicus* and *C. glacialis* is to examine their morphological characters (3) that have been suggested as diagnostic of the two species. Most common is the examination of the structure of the fifth pair of swimming legs in adult females and males (Jaschnov, 1955), and the morphology of the coxal endid of the mandible (gnathobase) (Beklemishev, 1959). Examination of both characters requires performing a fastidious and specific preparation on each specimen, and is therefore seldom applied during routine zooplankton samples analyses.

Although several diagnostic molecular markers have been developed for *Calanus*, from mtDNA RFLP (Lindeque et al., 1999) to nuclear InDels (Smolina et al., 2014), their use in the zooplankton research community has so far remained limited. A recent reappraisal of *Calanus* spp. distribution in the North Atlantic/Arctic Oceans relying on large scale sampling and molecular identification has suggested that misidentification is widespread and has led to erroneous conclusions regarding *Calanus* biogeography (M. Choquet et al. submitted).

Species misidentification may be less problematic in studies focusing on describing zooplankton assemblages based on higher taxonomic categories (e.g. Aßmus et al., 2009) or in trait-based studies, which aim at investigating ecological functions of assemblages (e.g. Brun et al., 2016). A correct species identification is however crucial for understanding species-specific life history strategies, species-specific productivity estimates and for studying distribution patterns, particularly if species are considered indicative for specific water masses and if changes in their distribution are assumed to have far reaching ecosystem impacts.

Both species differ in life strategies such as energy requirements for reproduction and growth, timing of reproduction, composition of overwintering populations and seasonal vertical migration patterns. These differences reflect adaptations to the environmental conditions in their main areas of distribution (Falk-Petersen et al., 2009), with *C. glacialis* having adapted more flexible life history strategy to deal with the constrains of seasonally ice-covered seas (Daase et al., 2013) and low temperature

leading to a larger body size and longer life span compared to *C. finmarchicus*. It is crucial to correctly identify them to understand their life history adaptations fully, how they have evolved differently in each species and how climate change will be affecting each species' productivity, population success, distribution and role in the food web. Using prosome length to discriminate between species has shown to underestimate smaller sized *C. glacialis* (Gabrielsen et al., 2012), which may bias species-specific biomass estimates and our understanding of energy allocations in that species.

In the present study, we use molecular tools to assess the reliability of the morphological characters used to discriminate between *C. finmarchicus* and *C. glacialis* across a large part of their distributional range.

Material and procedures

Samples collection and pre-sorting

Zooplankton were sampled in fjords along the Norwegian coast, in the White Sea, in Svalbard waters and in the Nansen Basin (Table 1) by vertically towed plankton nets (WP-2 / Juday types) with mesh sizes between 150 and 200 μm. The whole water column was sampled for most of the locations, except for the White Sea (100 – 0 m) and the Svalbard fjords (20 – 0 m). The sampling locations were selected to represent a latitudinal gradient from the southernmost (Lurefjord) to the northernmost (Nansen Basin) areas of co-occurring of *C. finmarchicus* and *C. glacialis*. The White Sea, where only *C. glacialis* occurrence was reported historically (Jaschnov, 1966, Jaschnov, 1955) and recently confirmed genetically (M. Choquet et al. submitted), and the region of Raunefjord / Korsfjord where only *C. finmarchicus* occurrence was reported, were also sampled in order to have more elements of comparison. Directly after sampling, a Folsom plankton splitter was used to randomly subsample ~100 to 200 live individuals of the older (≥CIV) copepodite stages. Prosome length measurements and examination of the redness of antennules and genital somite (for details see below) were carried out right after sampling, on the subsampled individuals kept alive in seawater. These live individuals were subsequently preserved individually in 70-80 % undenatured ethanol for later molecular-based species identification and morphological examinations.

Prosome length measurements

We subsampled up to 200 live individuals of late copepodite stages IV, V and CVI female (CIV, CV and CVIF) of *Calanus* per sample from each of the 9 locations (pooled into 6 geographically distant regions - 895 specimens in total) (Table 1). For the

sampling locations where it was possible, photographs of individuals were taken with a camera attached to a stereomicroscope. The prosome length of each specimen was measured from the tip of the cephalosome to the distal lateral end of the last thoracic somite (Fig. 1) either using the ruler in the eye-piece of a stereomicroscope to measure directly (resolution 1 µm), or by using cellSens Standard software (version 1.8.1 – Olympus corporation© 2009-2013) to analyse the photograph taken (resolution 0.01 µm). All the 895 individuals were identified with molecular markers (See section below "Molecular species identification"). Correlation between latitude and body size (prosome length) was tested independently for *C. finmarchicus* and *C. glacialis*, and separately for each developmental stage (CIV, CV, and CVI female) with use of Pearson´s correlation (in Microsoft® Excel® version 14.7.3).

Redness assessment

We evaluated the potential of red pigmentation ("redness") on antennules and genital somite to separate live *Calanus finmarchicus* and *C. glacialis,* as suggested in (Nielsen et al., 2014). A total of 903 *Calanus* individuals of developmental stage CIV to CVI (adult female and male) from 6 distant populations in the North Atlantic and Arctic Oceans were investigated in regard to their antennule redness (Table 1). Additionally, pigmentation of the two spermathecae on the ventral surface of the genital somite (the first urosome somite) in adult females was examined for 168 individuals from the same populations. All the individuals examined for their redness were subsequently identified with molecular markers (see section below "Molecular species identification").

The degree of antennules red pigmentation ("redness") was very heterogeneous among the studied individuals. We distinguished 4 different categories of individuals: antennules with more than 90% of redness; from 50 to 90% of redness; from 10 to 50% of redness; and less than 10% of redness. The percentage of redness used to distinguish different categories is based on the subjective evaluation of how much of the surface of antennules is red, and how dense this pigmentation is (see Fig. 2 for examples of each category). This choice is justified by our search for a parameter that could be easily and quickly used for routine species identification especially in the field.

Statistical differences in antennule redness between the two species *C. finmarchicus* and *C. glacialis*, among the different developmental stages, and among the locations sampled were tested using the Kruskal-Wallis H test.

To evaluate the pigmentation of the genital somite, we considered only 2 categories: red or pale (Fig. 3). Individuals with any redness on one or two of the spermathecae were assigned to "red"; the individuals for which no redness at all on the genital somite was noticeable were reported as "pale".

Molecular species identification

Each *Calanus* individual used for this study was genetically identified (913 individuals in total – Table 1). Molecular species identification followed the procedure described in M. Choquet et al. (submitted). In brief, DNA was extracted from animal's antennules using the HotSHOT DNA extraction method (Montero-Pau et al., 2008) and 6 nuclear molecular markers (Smolina et al., 2014) type InDels (Insertion or Deletion motifs) were amplified by Polymerase Chain Reaction (PCR). PCR amplicons were sized using a 3500xL Genetic Analyzer (Applied Biosystems, USA), generating a species specific profile (Smolina et al., 2014). Together, these 6 markers allow the reliable identification of *Calanus* species in the North Atlantic and Arctic Oceans (Nielsen et al., 2014, Smolina et al., 2014). This method allows to genotype each individual for species identification without using or destroying the animal's body. Once the antennules are removed, the rest of the body is intact and can still be examined for morphology.

5th pair of legs and gnathobase morphology examination

Seventy-one individuals from 5 different locations (Table 1) were examined (49 individuals of developmental stage CV and 22 individuals of CVI adult females), by following a specific procedure. M. Choquet selected the individuals among the genetically identified specimens preserved in ethanol, in order to have both species represented. The 71 selected ones were sent to S. Kwaśniewski for dissection (see procedure below), without giving any information about the molecular results of species ID for these particular individuals. After dissection, photographs of the dissected body parts were taken for each individual by S. Kwaśniewski, and shared with K. Kosobokova. Examination of the fifth thoracic leg (swimming leg P5 – Fig. 4 a-b & Fig. 5 a-b) and the coxal endid of the mandible (gnathobase – Fig. 4 c-d & Fig. 5 c-d) were carried out by both S. Kwaśniewski and K. Kosobokova independently, based on the photographs only. Their species identification decisions, based on the pictures analysis, were then sent back to M. Choquet to compare with molecular results. We decided to follow this approach in order to avoid any bias in the expert interpretation of the pictures due to the prior knowledge of molecular ID.

For the examination of the P5 morphology, descriptions of the leg structure provided in Jaschnov (1955); Frost (1974); and Jaschnov, (1955), Frost, (1974), Brodskii et al., (1983) were used. The P5 in *Calanus* consists of a remnant of precoxa, well developed coxa (basipod 1) and basis (basipod 2), from which two 3-segmented rami (exopod and endopod) grow out (Huys and Boxshall, 1991). Investigation focused on the lamellar structure with denticulated edge, the denticulated lamella, extending longitudinally on the medial margin of the coxa of P5 from the intercoxal plate to near the distal medial corner, little folding on the posterior surface of the basipod segment. According to the references, in *C. finmarchicus* the denticulated lamella is straight,

missing clearly expressed incurvation characteristic for *C. glacialis* (See Fig. 4 a-b). In *C. glacialis*, the denticulated lamella is concaved and has a well-expressed curvature (deflection) (Frost, 1974, Jaschnov, 1955) (See Fig. 5 a-b).

For examination of the gnathobase, descriptions provided in (Beklemishev, 1959, Vyshkvartzeva, 1976, Vyshkvartzeva, 1972) were used. The gnathobase is the coxal endite (a medially directed process on the protopodal segment of the appendage), bearing the toothed cutting edge distally (Huys and Boxshall, 1991). The cutting (masticatory) edge of the gnathobase bears several groups of teeth varying in form and structure. Some of these teeth (at least in sexually developed stages) are covered with silicate crowns. In adult females of *C. glacialis* and *C. finmarchicus* the complete arrangement of gnathobase cutting edge includes ventral (V_1 – V_2), central (C_1 – C_4) and distal (D_1 – D_3) teeth plus flexible setae with one or two rows of spines. Between groups of V and C teeth there is a diastema (a gap between the teeth). Tooth V_2 does not have a crown and teeth of group D are often equipped on their lateral surfaces with small denticles. According to Beklemishev (1959) and Vyshkvartzeva (1976, 1972), species-specific differences in the form and arrangement of the teeth concern teeth V_1 and V_2. In *C. glacialis* adult females, the crown of the tooth V_1 is not very high, compressed in the anterior-posterior direction, and has 2-3 peaks. The tooth V_2, which does not have a crown, is well-developed and placed on wide cuticular platform. Its size is close to the size of V_1 and it approximately equals to the diameter of its base (Fig. 5 c-d). In *C. finmarchicus* adult females, the tooth V_2 is smaller than V_1 and its height is larger than the diameter of its base, but its form and size varies (Fig. 4 c-d). In comparison with *C. glacialis*, the tooth V_2 in *C. finmarchicus* presents as not completely formed.

The examination of the two structures was done after dissection and slides preparation. Each individual from the study collection was first immersed for 10 min in a drop of glycerol: ethanol 1:1 mixture placed on a microscope slide with cavity. In 10 min each individual was photographed using Olympus SC50 CMOS Color Camera, mounted with a photo adapter U-TV0.5xc-3 on Olympus SZX12 Research Stereomicroscope, equipped with AXH1x and DFPL2x-3 objectives. The acquisition of the digital pictures was made with Olympus cellSense Imaging Software v.1.12. The pictures of the body habitus of each individual were made at 10x total magnification, one picture with use of AXH1x objective and one with use of DFPL2x-3 objective. Then the two structures under consideration were dissected from the body. The P5 was cut off the thoracic somite and placed in a drop of the same glycerol: ethanol 1:1 mixture, on a regular microscope slide, anterior side upward.

The mandibles were also dissected one by one from the cephalosome. After removal of the mandible, the gnathobase was dissected from the appendage, and mounted in

another drop of glycerol: ethanol 1:1 mixture, anterior side upward. The same procedure was repeated for the second mandible, and finally the pair of gnathobases belonging to one individual was covered with a glass coverslip. The dissection of the appendages and preparation of the microscope slides was done with use of Olympus SZX12 stereomicroscope, at magnifications ranging from 7x to 90x. In the following step, the investigated structures were photographed using Olympus SC30 CMOS Color Camera, mounted with a photo adapter U-TV1x-2 on Olympus BX51 system microscope, equipped with PlanN 4x and UPlanFLN 10x objectives. The acquisition of the digital pictures was made with use of Olympus cellB Imaging Software v.3.3.

Assessment

Prosome length measurements

Based on prosome length measurements of 895 genetically-identified individuals from six regions, we confirm that this character shows a global overlap of size between *C. finmarchicus* and *C. glacialis* regardless of developmental stage (Fig. 6). Size frequency distributions, however, differed among different regions. In the Norwegian fjords (Saltenfjord / Skjerstadfjord; Lurefjord), *C. glacialis* showed a complete size overlap with *C. finmarchicus*, but these *C. glacialis* were significantly smaller (t-test, $p < 0.01$) than the *C. glacialis* captured in the White Sea and the high Arctic. Noteworthy, our data showed positive correlations between latitude and body size for both *C. glacialis* and *C. finmarchicus* (Table 2).

Thus, the prosome length cannot reliably discriminate between *C. finmarchicus* and *C. glacialis* in any of the investigated regions, and even less in the Norwegian fjords. However, in the Nansen Basin and Svalbard waters, the majority of the length values for *C. finmarchicus* and *C. glacialis* follow a dichotomy. Prosome length could therefore be used in those particular areas to approximate the overall *C. glacialis* and *C. finmarchicus* composition. It has to be kept in mind, however, the inaccuracy of the method leading to underestimation of *C. glacialis* (especially small-sized individuals), and over-estimation of *C. finmarchicus* numbers (Gabrielsen et al., 2012).

Redness assessment

We tested if redness can be used to reliably separate between live *C. finmarchicus* and *C. glacialis*, at different developmental stages and across different regions of co-occurrence. According to Nielsen et al. (2014), the genital somite (originally genital field) and the antennules of *C. glacialis* adult females had red pigmentation, while the pigmentation of female *C. finmarchicus* were mostly pale. However, the study focused

only on adult females from a limited geographic location (Greenland) (Nielsen et al., 2014).

In our study, red pigmentation of antennules was variable (Fig. 7-8, Supplementary 2a-b), with significant differences in redness between the two species, among developmental stages, and locations sampled (Supplementary 1). Antennules redness was assessed for 903 individuals, from copepodite stage CIV to adult females and males, at six different locations (Table 3a-3f & Fig. 7). Molecular identification of these 903 individuals was performed consecutively.

At the northernmost location, the Nansen Basin (Table 3a), all genetically identified *C. glacialis* had >10% redness on their antennules, and they were all adult females. Stages CV and adult females of genetically identified *C. finmarchicus* individuals collected in the same place were mainly pale except for 3 females with a slight redness (10 to 50%).

In Svalbard, the majority of *C. glacialis* identified genetically, including stages CIV, CV and adult females, had also >10% redness (except 1 CV and 1 adult female with <10% – Table 3b). *C. finmarchicus* individuals, including stages CIV, CV and CVI, tended to be paler compared to *C. glacialis* in Svalbard, but some *C. finmarchicus*, especially females, had >10% redness. One male was detected there, identified as *C. finmarchicus* with pale antennules.

Only *C. glacialis* was detected in the White Sea sample (stages CIV, CV and adult females - Table 3c). Individuals from stage CIV exhibited almost none, or very little (less than 50%) redness, but a stronger red pigmentation was observed for the older stages CV and adult females.

In the boreal fjords Saltenfjord and Skjerstadfjord (Table 3d), *C. glacialis* individuals (stages CIV to CVI) most often (88%) had red pigments. All males, 4% of the females and 20% of the CV *C. glacialis* were pale. The majority of *C. finmarchicus* individuals were pale in these two fjords, independently of the developmental stage, however, with 3 exceptions (1 CIV and 2 adult females). Interestingly, males of both species were totally pale.

In Lurefjord (southern Norway – Table 3e), the majority (73%) of *C. glacialis* had red pigmentation, with 25% pale CV and 100% pale females. In comparison, the majority (84%) of *C. finmarchicus* (CVs) were pale there.

In the open southern fjords Raunefjord and Korsfjord (Table 3f), we only identified *C. finmarchicus* among the older stages (CIV and CV) in our samples and 56% of these were pale and another 16% had 10-50% redness.

Despite identifying significant differences (Kruskal-Wallis H test) in the redness of antennules between species, among stages for each species, and among locations for each species for every set of variables compared (Supplementary 1), the general trend was that the majority of individuals of *C. finmarchicus* tends to have pale antennas whereas the majority of *C. glacialis* tends to have red ones (Fig. 7 & 8). In both species, there were exceptions, especially for *C. glacialis* in the White Sea and *C. finmarchicus* in Raunefjord / Korsfjord (Supplementary 2a). The tendencies in pigmentation were similar for the different developmental stages (Supplementary 2b) except that males of both species were pale without exception (albeit only a few males were investigated) and *C. glacialis* CIV in general being less pigmented than *C. glacialis* CV and adult females. Antennules redness thus appears not to be a reliable diagnostic feature and is clearly not a species-specific trait. It was never 100% diagnostic for any of the 6 regions investigated. Assessment of pigmentation might be useful to get an overall impression of the species composition in the Arctic Ocean and in isolated fjords, taking into account the error threshold (region dependent), and the fact that investigations have to be done on live organisms.

Regarding the redness of the genital somite of *Calanus* females, all the *C. finmarchicus* examined had pale spermathecae, although we only found females of this species in Svalbard and Saltenfjord / Skjerstadfjord (Table 4). Most of the *C. glacialis* examined (from 4 regions) had red genital somite, but also a few individuals had pale genital somite in each region. Our results indicate that redness of genital somite is also not 100% diagnostic for species identification. However, the character seems to be useful to get a global idea of species composition of a zooplankton sample, but using it may result in an underestimation of *C. glacialis* number of individuals.

5th pair of legs and gnathobase morphology examination

The curvature of the inner denticulated margin of P5 swimming legs and the shape of the mandibular cutting blade are morphological characters that have been described early in the literature as species-specific (Beklemishev, 1959, Brodskii et al., 1983, Frost, 1974, Jaschnov, 1955, Vyshkvartzeva, 1976). However, due to the arduousness of their examination, they remain rarely used to identify *Calanus* species.

Only 23 individuals out of the 71 examined exhibited the species-specific features typical for the species they belong to (verified by genetics), according to the literature (Supplementary 3). For the other individuals, the morphological characteristics examined were different from that of the species according to the literature (Fig. 9 & 10**)**. Furthermore, no geographic coherence was found in the deviations of the

characteristics (Supplementary 3). This resulted in an error rate of 30% and 31% in the identification decisions made by the experts in *Calanus* morphology, after comparing their decision with results of genetic identification. Identification decisions of both experts matched only 36 times, and of these only 32 individuals (45% of the total) were confirmed to be correct by genetic identification. More specifically, experts' decision and genetics matched at 51% for the individuals at stage CV, while experts' decision and genetics only matched at 32% for the adult females individuals. It has to be kept in mind that the morphological features described in literature to discriminate between *Calanus* species are typically described and can be applied directly for identification of adult females (or males) only, while we tested them on both adult females and CVs. They may not work for distinguishing copepodids at pre-adult CV stage, as some morphological structures are still not fully developed or expressed. However, the misidentification of 68% of adult females and disagreement between two experts is striking. In a few cases, the characteristics observed in genetically identified species had appearance theoretically typical of the opposite species. Part of the problem may result from the fact that the characteristics are at the moment predominantly of a descriptive type and they have been portrayed based on "typical" individuals from only a few sites over the species distribution range.

To conclude, the morphological characters involving the 5th pair of legs and the gnathobase were not consistent enough to be used for species identification. Therefore, we cannot recommend using these characteristics to reliably identify *C. finmarchicus* and *C. glacialis* without additional investigations.

Discussion

Characters variability

The smaller size of *C. glacialis* in the Norwegian fjord populations, compared to high Arctic populations, largely explains why the species wide boreal occurrence (M. Choquet et al. submitted) has not been detected before. For instance, occurrence of *C. glacialis* in the southern Lurefjord was not detected before molecular markers were applied (Bucklin et al., 2000). In the context of climate change and ocean warming, it is to be expected that more and more *C. glacialis* individuals will be able to complete their life-cycle within a year, and then have a body size comparable to that of *C. finmarchicus.* The decrease in body size with decreasing latitude is likely a direct effect of temperature (Atkinson and Sibly, 1997), but variation in the duration of the productive season and predation pressure by visual predators (Brooks and Dodson, 1965) may also play an important role. Copepods are ectothermic, they primarily rely on external sources to regulate their body heat. The temperature-size-rule (TSR) refers

to the widely observed phenomenon that ectotherms reared at lower temperatures usually grow more slowly, but become larger as adults compared to individuals reared at higher temperatures (Atkinson and Sibly, 1997, Atkinson, 1994). Calanoid copepods appear especially sensitive to temperature by having a four-fold greater reduction in adult body mass per degree Celsius compared to Cyclopoid copepods (Horne et al., 2016). Increasing latitude and mean temperature are strongly correlated (Sunday et al., 2011), and distinguishing separate effects may not be straight forward. However, oxygen demand and supply has been suggested as a driver of both processes (Horne et al., 2015), as the metabolic demand increases with increasing temperature, while the oxygen availability in the water decreases (Verberk et al., 2011). In addition, on-going climate change that is impacting the temperature of *Calanus* habitat brings another unpredictable variable affecting body size of *Calanus* species. Predation by visual predators, such as fish, may also induce a change in body-size composition in zooplankton communities. In the classical study by (Brooks and Dodson, 1965) , the zooplankton community shifted from dominance of large- to dominance of small species in a freshwater lake after a fish-predator was introduced. According to optimal foraging theory, predators should target larger sized prey when handling time is a restriction. Both modelling studies and field investigation confirm that lesser sandeel (*Ammodytes marinus*) in the North Sea actively target large copepods, such as *C. finmarchicus,* over smaller copepod taxa when these are available (van Deurs et al., 2015, van Deurs et al., 2014). On a longer time-scale, adaptive responses to predation pressure on the larger species may result in a dominance of species with shorter life-spans and smaller body-size (Berge et al., 2012, Stearns, 1992). However, to the best of our knowledge there are no studies showing that predation may cause intraspecific changes in body size within populations of *Calanus* spp..

It has been proposed that the pigment involved in redness of Calanoid copepods is astaxanthin, a form of keto-carotenoid (Mojib et al., 2014). This pigment has a role in the protection against UVR irradiance, and usually appears red in copepods. Copepods can adjust their level of astaxanthin pigment quickly, even within a season, depending on the prevailing threat, UVR or predators (Hansson, 2000). Given such variability it is thus not surprising that redness cannot be used reliably as a species diagnostic tool. Examination of more samples for each developmental stage, from different depths, and seasonal observations may help to better understand the reasons for variability of red pigmentation in *Calanus* and its relation to environmental parameters.

Biological implications

Copepod species of the genus *Calanus* are the most studied amongst the zooplankton. They are often used as biological indicators of water masses and to follow the effects of climate change on the marine ecosystems. However, in the

majority of past studies, species identification has been based on morphometric and morphological characteristics. We found that none of the morphometric and morphological characteristics used in literature allow for unequivocal identification and separation of species. Therefore, it is likely that our knowledge of *Calanus* geographical distribution is plagued by species misidentification. Indeed, a large part of the distribution range of *C. glacialis* has only been recently identified along the Norwegian coast (M. Choquet et al. submitted), questioning the necessity of direct connection of its populations to Arctic waters. In other words, life cycles, phenology and exact role of each species within fjord ecosystems, potential for adaptability/resilience to climate variability, as well as response to environmental variations and population dynamics are not fully understood. Predictions on climate change effects and ecological models based on the present view of *Calanus* distribution and stocks dynamics are likely to be at least partially erroneous, especially in the areas of sympatry.

Comments and recommendations

None of the morphological characters described in literature and re-assessed in the present study can reliably identify *Calanus finmarchicus* and *C. glacialis* with 100% confidence. There are some global trends that can bring information about the species composition though, but certainly not equally everywhere. Prosome length may be useful to approximate the species composition in the Nansen Basin and in Svalbard waters, and likely in the Arctic Ocean. However, it is critical to keep in mind the underestimation of *C. glacialis*. In fjords along the Norwegian coast, prosome length is clearly not usable, as the size range of both species overlaps completely. Regarding the redness of antennules / genital somite of *Calanus*, it seems to be a useful indicator of species in the Arctic and in relatively closed fjords (with a sill - e.g. Saltenfjord, Skjerstadfjord, Lurefjord), but not in open fjords (without sill – e.g. Raunefjord, Korsfjord). Again, by using this character it is critical to keep in mind the variable error rates associated (Fig. 7). We recommend not using the curvature of the inner denticulated margin of the P5 swimming legs and the shape of the mandibular cutting blade to discriminate between species, until the variability of these characters in all parts of the species distribution range is thoroughly investigated simultaneously with molecular identification.

The use of molecular tools is thus the only reliable method for discriminating between the two species. It is likely that the problems of identification encountered with *Calanus* also exist in other taxa in pelagic zooplankton (e.g. Aarbakke et al., 2011). Therefore, it is critical to start using molecular tools routinely for reliable species

identification, especially for ecologically important organisms such as *Calanus*. Equipment, time, competences needed and cost related to molecular identification of *Calanus* are today a much lesser issue than it used to be. Indeed, as described in (Smolina et al., 2014), the set of InDels markers that we used in the present study can be ran on agarose gels and therefore used in a low-cost setting on board a research vessel. We also simplified the method of DNA extraction, which now consists in only removing the antennules of each individuals and incubating them 30 minutes in a buffer at no costs. With these simplifications, genotyping 96 individuals of *Calanus* can be done in 5 hours for less than 2 USD per individual.

Acknowledgements

We are very grateful to Morten Krogstad and to the crew and captain of the R/V G.O. Sars for their help with the sampling. We thank Mikko Vihtakari for his suggestions and practical help on treatment of pigmentation data and statistics. The project was funded by the European Commission FP7 EURO-BASIN (Grant agreement: 264 933), the Norwegian Research Council (project HAVKYST 216578 and PolarProg 227139 and 246747) and Nord University. M.C. was supported by a PhD student fellowship from Nord University, and M.H. by a PhD student fellowship from The University Centre in Svalbard. K.K. was supported by Russian Foundation for Basic Research (15-29-02447; 16-04-00375) and Russian Scientific Foundation (14-50-00095). M.D. was supported by NRC grant 226417 Marine Night. S.K. was supported by the Polish-Norwegian Research Program Pol-Nor/201992/93/2014 (Project DWARF). We also thank the research network ARCTOS for useful collaborations.

References

Aarbakke, O.N.S., Bucklin, A., Halsband, C. & Norrbin, F. (2011). Discovery of *Pseudocalanus moultoni* (Frost, 1989) in Northeast Atlantic waters based on mitochondrial COI sequence variation. *Journal of Plankton Research,* 33**:** 1487-1495.

Arnkværn, G., Daase, M. & Eiane, K. (2005). Dynamics of coexisting *Calanus finmarchicus*, *Calanus glacialis* and *Calanus hyperboreus* populations in a high-Arctic fjord. *Polar Biology,* 28**:** 528-538. DOI 10.1007/s00300-005-0715-8.

Aßmus, J., Melle, W., Tjøstheim, D. & Edwards, M. (2009). Seasonal cycles and long-term trends of plankton in shelf and oceanic habitats of the Norwegian Sea in relation to environmental variables. *Deep Sea Research Part II: Topical Studies in Oceanography,* 56**:** 1895-1909.

Atkinson, D. & Sibly, R.M. (1997). Why are organisms usually bigger in colder environments? Making sense of a life history puzzle. *Trends in Ecology & Evolution,* 12**:** 235-239. DOI 10.1016/S0169-5347(97)01058-6.

Atkinson, D. (1994). Temperature and Organism Size - a Biological Law for Ectotherms. *Advances in Ecological Research, Vol 25.*

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S. & Reid, P.C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature,* 426**:** 661-664.

Beklemishev, K. (1959). On the anatomy of masticatory organs of Copepoda. Report 2: The masticatory edge in mandibles of certain species of Calanidae and Eucalanidae. *Trudy Instituta Okeanologii,* 30**:** 148-155.

Berge, J., Gabrielsen, T.M., Moline, M. & Renaud, P.E. (2012). Evolution of the Arctic *Calanus* complex: an Arctic marine avocado? *J Plankton Res,* 34**:** 191-195. DOI 10.1093/plankt/fbr103.

Blachowiak-Samolyk, K. (2008). Contrasting zooplankton communities (Arctic vs. Atlantic) in the European Arctic marginal ice zone. *Oceanologia,* 50**:** 363-389.

Bonnet, D. & Frid, C. (2004). Seven copepod species considered as indicators of water-mass influence and changes: results from a Northumberland coastal station. *ICES Journal of Marine Science,* 61**:** 485-491.

Brodskii, K., Vyshkvartseva, N., Kos, M. & Markhatseva, E. (1983). Copepods (Copepoda: Calanoida) of the seas of the USSR and adjacent waters. *Keys to the fauna of the USSR,* 1**:** 358.

Broms, C., Melle, W. & Kaartvedt, S. (2009). Oceanic distribution and life cycle of *Calanus* species in the Norwegian Sea and adjacent waters. *Deep Sea Research Part II: Topical Studies in Oceanography,* 56**:** 1910-1921. DOI 10.1016/j.dsr2.2008.11.005.

Brooks, J.L. & Dodson, S.I. (1965). Predation, body size, and composition of plankton. *Science,* 150**:** 28-35.

Brun, P., Payne, M.R. & Kiørboe, T. (2016). Trait biogeography of marine copepods – an analysis across scales. *Ecology Letters***:** n/a-n/a. DOI 10.1111/ele.12688.

Bucklin, A., Kaartvedt, S., Guarnieri, M. & Goswami, U. (2000). Population genetics of drifting (*Calanus* spp.) and resident (*Acartia clausi*) plankton in Norwegian fjords. *Journal of Plankton Research,* 22**:** 1237-1251.

Conover, R. (1988). Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. *Biology of Copepods.* Springer.

Daase, M., Falk-Petersen, S., Varpe, Ø., Darnis, G., Søreide, J.E., Wold, A., Leu, E., Berge, J., Philippe, B. & Fortier, L. (2013). Timing of reproductive events in the marine copepod *Calanus glacialis*: a pan-Arctic perspective. *Canadian journal of fisheries and aquatic sciences,* 70**:** 871-884.

Daase, M. & Eiane, K. (2007). Mesozooplankton distribution in northern Svalbard waters in relation to hydrography. *Polar Biology,* 30**:** 969-981.

Falk-Petersen, S., Mayzaud, P., Kattner, G. & Sargent, J. (2009). Lipids and life strategy of Arctic *Calanus*. *Marine Biology Research,* 5**:** 18-39. DOI 10.1080/17451000802512267.

Falk-Petersen, S., Pavlov, V., Timofeev, S. & Sargent, J.R. (2007). Climate variability and possible effects on arctic food chains: the role of *Calanus*. *Arctic alpine ecosystems and people in a changing environment Springer, Berlin***:** 147-166.

Fleminger, A. & Hulsemann, K. (1977). Geographical range and taxonomic divergence in North Atlantic *Calanus* (*C. helgolandicus*, *C. finmarchicus* and *C. glacialis*). *Marine Biology,* 40**:** 233-248.

Forest, A., Tremblay, J.-É., Gratton, Y., Martin, J., Gagnon, J., Darnis, G., Sampei, M., Fortier, L., Ardyna, M. & Gosselin, M. (2011). Biogenic carbon flows through the planktonic food web of the Amundsen Gulf (Arctic Ocean): A synthesis of field measurements and inverse modeling analyses. *Progress in Oceanography,* 91**:** 410-436.

Frost, B. (1974). *Calanus marshallae*, a new species of calanoid copepod closely allied to the sibling species *C. finmarchicus* and *C. glacialis*. *Marine Biology,* 26**:** 77-99.

Gabrielsen, T.M., Merkel, B., Søreide, J.E., Johansson-Karlsson, E., Bailey, A., Vogedes, D., Nygård, H., Varpe, Ø. & Berge, J. (2012). Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. *Polar Biology,* 35**:** 1621-1628. DOI 10.1007/s00300-012-1202-7.

Gislason, A. & Astthorsson, O.S. (2002). The food of Norwegian spring-spawning herring in the western Norwegian Sea in relation to the annual cycle of zooplankton. *Sarsia: North Atlantic Marine Science,* 87**:** 236-247.

Hansson, L.A. (2000). Induced pigmentation in zooplankton: a trade–off between threats from predation and ultraviolet radiation. *Proceedings of the Royal Society of London B: Biological Sciences,* 267**:** 2327-2331.

Hassel, A. (1986). Seasonal changes in zooplankton composition in the Barents Sea, with special attention to *Calanus* spp.(Copepoda). *Journal of Plankton Research,* 8**:** 329-339.

Hays, G.C., Richardson, A.J. & Robinson, C. (2005). Climate change and marine plankton. *Trends in ecology & evolution,* 20**:** 337-344.

Helaouët, P. & Beaugrand, G. (2007). Macroecology of *Calanus finmarchicus* and *C. helgolandicus* in the North Atlantic Ocean and adjacent seas. *Marine Ecology Progress Series,* 345**:** 147-165.

Hirche, H.-J. & Kosobokova, K. (2011). Winter studies on zooplankton in Arctic seas: the Storfjord (Svalbard) and adjacent ice-covered Barents Sea. *Marine biology,* 158**:** 2359.

Hirche, H.-J., Hagen, W., Mumm, N. & Richter, C. (1994). The Northeast Water Polynya, Greenland Sea. *Polar Biology,* 14**:** 491-503.

Horne, C.R., Hirst, A.G., Atkinson, D., Neves, A. & Kiørboe, T. (2016). A global synthesis of seasonal temperature-size responses in copepods. *Global Ecology and Biogeography,* 25**:** 988-999. DOI 10.1111/geb.12460.

Horne, C.R., Hirst, A.G. & Atkinson, D. (2015). Temperature-size responses match latitudinal-size clines in arthropods, revealing critical differences between aquatic and terrestrial species. *Ecology Letters,* 18**:** 327-335. DOI 10.1111/ele.12413.

Huys, R. & Boxshall, G.A. (1991). *Copepod evolution*: Ray Society.

Jaschnov, W. (1972). On the systematic status of *Calanus glacialis*, *Calanus finmarchicus* and *Calanus helgolandicus*. *Crustaceana,* 22**:** 279-284.

Jaschnov, W. (1970). Distribution of *Calanus* species in the seas of the northern hemisphere. *Internationale Revue der gesamten Hydrobiologie und Hydrographie,* 55**:** 197-212.

Jaschnov, W. (1955). Morphology, distribution and systematics of *Calanus finmarchicus* sl. *Zool Zh,* 34**:** 1210-1223.

Jaschnov, W. (1966). Water mass and Plankton: 4. *Calanus finmarchicus***:** 404-412.

Kosobokova, K. (2012). Zooplankton of the Arctic Ocean: Community structure, Ecology, Spatial distribution. *GEOS, Moscow, 272p (in Russian)*.

Kosobokova, K.N., Hopcroft, R.R. & Hirche, H.-J. (2011). Patterns of zooplankton diversity through the depths of the Arctic's central basins. *Marine Biodiversity,* 41**:** 29-50.

Kosobokova, K. & Hirche, H.-J. (2009). Biomass of zooplankton in the eastern Arctic Ocean–a base line study. *Progress in Oceanography,* 82**:** 265-280.

Kwasniewski, S., Hop, H., Falk-Petersen, S. & Pedersen, G. (2003). Distribution of *Calanus* species in Kongsfjorden, a glacial fjord in Svalbard. *Journal of plankton research,* 25**:** 1-20.

Lindeque, P.K., Hay, S.J., Heath, M.R., Ingvarsdottir, A., Rasmussen, J., Smerdon, G.R. & Waniek, J.J. (2006). Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. *Journal of Plankton Research,* 28**:** 221-238. DOI 10.1093/plankt/fbi115.

Lindeque, P., Harris, R., Jones, M. & Smerdon, G. (1999). Simple molecular method to distinguish the identity of *Calanus* species (Copepoda: Calanoida) at any developmental stage. *Marine Biology,* 133**:** 91-96.

Michaud, J. & Taggart, C.T. (2007). Lipid and gross energy content of North Atlantic right whale food, *Calanus finmarchicus*, in the Bay of Fundy. *Endangered Species Research*.

Mojib, N., Amad, M., Thimma, M., Aldanondo, N., Kumaran, M. & Irigoien, X. (2014). Carotenoid metabolic profiling and transcriptome-genome mining reveal functional equivalence among blue-pigmented copepods and appendicularia. *Molecular ecology,* 23**:** 2740-2756.

Montero-Pau, J., Gómez, A. & Muñoz, J. (2008). Application of an inexpensive and high-throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods,* 6**:** 218-222.

Nielsen, T.G., Kjellerup, S., Smolina, I., Hoarau, G. & Lindeque, P. (2014). Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phenological differences? *Marine Biology,* 161**:** 1299-1306. DOI 10.1007/s00227-014-2419-5.

Parent, G.J., Plourde, S. & Turgeon, J. (2011). Overlapping size ranges of *Calanus* spp. off the Canadian Arctic and Atlantic Coasts: impact on species' abundances. *Journal of Plankton Research,* 33**:** 1654-1665. DOI 10.1093/plankt/fbr072.

Skjoldal, H.R. (2004). An introduction to the Norwegian Sea ecosystem. *The Norwegian Sea Ecosystem Tapir Academic Press, Trondheim***:** 15-32.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14**:** 1072-9. DOI 10.1111/1755-0998.12241.

Søreide, J.E., Falk-Petersen, S., Hegseth, E.N., Hop, H., Carroll, M.L., Hobson, K.A. & Blachowiak-Samolyk, K. (2008). Seasonal feeding strategies of *Calanus* in the high-Arctic Svalbard region. *Deep Sea Research Part II: Topical Studies in Oceanography,* 55**:** 2225-2244.

Stearns, S.C. (1992). *The Evolution of Life Histories* Oxford: Oxford University Press.

Steen, H., Vogedes, D., Broms, F., Falk-Petersen, S. & Berge, J. (2007). Little auks (*Alle alle*) breeding in a High Arctic fjord system: bimodal foraging strategies as a response to poor food quality? *Polar Research,* 26**:** 118-125.

Sunday, J.M., Bates, A.E. & Dulvy, N.K. (2011). Global analysis of thermal tolerance and latitude in ectotherms. *Proceedings of the Royal Society B-Biological Sciences,* 278**:** 1823-1830. DOI 10.1098/rspb.2010.1295.

Unstad, K.H. & Tande, K.S. (1991). Depth distribution of *Calanus finmarchicus* and *C. glacialis* in relation to environmental conditions in the Barents Sea. *Polar Research,* 10**:** 409-420.

Van Deurs, M., Jorgensen, C. & Fiksen, O. (2015). Effects of copepod size on fish growth: a model based on data for North Sea sandeel. *Mar Ecol Prog Ser,* 520**:** 235-243. DOI 10.3354/meps11092.

Van Deurs, M., Koski, M. & Rindorf, A. (2014). Does copepod size determine food consumption of particulate feeding fish? *Ices Journal of Marine Science,* 71**:** 35-43. DOI 10.1093/icesjms/fst090.

Varpe, Ø., Fiksen, Ø. & Slotte, A. (2005). Meta-ecosystems and biological energy transport from ocean to coast: the ecological importance of herring migration. *Oecologia,* 146**:** 443.

Verberk, W., Bilton, D.T., Calosi, P. & Spicer, J.I. (2011). Oxygen supply in aquatic ectotherms: Partial pressure and solubility together explain biodiversity and size patterns. *Ecology,* 92**:** 1565-1572.

Vyshkvartzeva, N. (1976). The functional morphology of mouth parts of the species *Calanus* sl (Copepoda, Calanoida). *Essled Fauny Morey,* 18**:** 11-69.

Vyshkvartzeva, N. (1972). The structure of mandibles in Copepoda (the genus *Calanus*) in relation to the latitudinal zonality. *Issled Fauny Morey,* 12**:** 161-171.

Wassmann, P., Kosobokova, K., Slagstad, D., Drinkwater, K., Hopcroft, R., Moore, S., Ellingsen, I., Nelson, R., Carmack, E. & Popova, E. (2015). The contiguous domains of Arctic Ocean advection: trails of life and death. *Progress in Oceanography,* 139**:** 42-65.

Wassmann, P., Reigstad, M., Haug, T., Rudels, B., Carroll, M.L., Hop, H., Gabrielsen, G.W., Falk-Petersen, S., Denisenko, S.G. & Arashkevich, E. (2006). Food webs and carbon flux in the Barents Sea. *Progress in Oceanography,* 71**:** 232-287.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis Springer-Verlag. *New York*.

Wilson, S.E., Steinberg, D.K. & Buesseler, K.O. (2008). Changes in fecal pellet characteristics with depth as indicators of zooplankton repackaging of particles in the mesopelagic zone of the subtropical and subarctic North Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography,* 55**:** 1636-1647.

Figures legends

Figure 1: ***Calanus finmarchicus*, structure of the body**.

Figure 2: Categories of red pigmentation of antennules in *Calanus*. Two photos are shown as examples for each of the four categories defined as follows: less than 10% of red pigmentation; between 10 and 50% pigmentation; between 50 and 90% of pigmentation; and more than 90% pigmentation.

Figure 3: Pigmentation of genital somite in *Calanus*. Pigmentation is defined as red (left photo) or pale (right photo).

Figure 4: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus finmarchicus* (genetically confirmed) exhibiting the species-specific traits as described in literature. Specimen collected from Van Mijenfjord (ID: VM41). **a** & **b.** Anterior view of the fifth thoracic leg (P5) with denticulated lamellae on the medial margin of the coxa, showing typical "straight form". Abbreviations used: coxa (c); basis (b); - exopods 1 to 3 (ex1-ex3); - endopods 1 to 3 (en1-en3). **c** & **d**. Anterior view of mandible gnathobases, with a typical small second ventral tooth on the cutting edge. Abbreviations: left mandible gnathobase (gnth le); right mandible gnathobase (gnth ri); first ventral tooth (V1); second ventral tooth (V2).

Figure 5: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus glacialis* (genetically confirmed) exhibiting the species-specific traits as described in literature. Specimen collected from Skjerstadjord (ID: SKJ25). **a** & **b.** Anterior view of the 5th thoracic leg (P5) showing denticulated lamellae on the medial margin of the coxa in a typical concave form, with well-expressed curvature. **c** & **d**. Anterior view of mandible gnathobases with the cutting edge with a typical large second ventral tooth on a wide basis (See legend Fig. 4 for abbreviations meaning).

Figure 6: Stage-specific length frequency distributions of prosome length (mm) for copepodites CIV, CV and adult females of *Calanus glacialis* and *C. finmarchicus* in different regions. In total, 895 individuals were measured, from 9 locations, pooled into 6 distant regions, in the North Atlantic and Arctic Oceans. Only *C. glacialis* (blue) occurred in the White Sea, and only *C. finmarchicus* (red) occurred in Raunefjord / Korsfjord area.

Figure 7: Red pigmentation on *Calanus finmarchicus* (*C. fin*) and *C. glacialis* (*C. gla*) antennules in different regions. Species-related redness from 4 regions where both species co-occur: the Nansen Basin, Svalbard, Saltenfjord / Skjerstadfjord and Lurefjord; and from the White Sea where only *C. glacialis* occurs, and Raunefjord / Korsfjord where only *C. finmarchicus* occurs. Blue colour of the pie charts indicates proportion of individuals for which less than 10% of the surface of their antennules was red; red colour indicates proportion of individuals for which more than 10% of red pigmentation was noticed.

Figure 8: Antennules redness frequency distribution per *Calanus* species. This violin graph was realised under RStudio v.1.0.143 with the package ggplot2 (Wickham, 2009). The graph shows the distribution of each species individuals on the following 3 ranks scale of redness: 1 = less than 10% of redness; 2 = 10-50% redness; 3 = more than 50% redness.

Figure 9: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus finmarchicus* exhibiting traits theoretically assigned to *C. glacialis*. The specimen from Saltenfjord (ID: SALT27) exhibits concave denticulated lamellae with a well-expressed curvature, and a wide basis of the second ventral tooth, typical of *C. glacialis* according to literature.

Figure 10: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus glacialis* exhibiting traits theoretically assigned to *C. finmarchicus*. The specimen from Saltenfjord (ID: SALT14) exhibits straight shaped denticulated lamellae, and a small second ventral tooth on the cutting edge of the coxa, typical of *C. finmarchicus* according to literature.

Table 1: Sampling locations with positions, sampling dates and number of individuals used for each analysis.

Location		GPS	Date	N ind. analysed			
				PL	Redness		Legs / Gnath
					Ant	Gen	
Nansen Basin		87°00 N 55°47 E	10/4/16	96	94	0	0
Svalbard	Isfjord	78°19 N 15°09 E	6/5/16	136	227	60	0
	Van Mijenfj. (VM)	77°46 N 15°02 E	6/3/16	90			16
White Sea		66°33 N 33°43 E	8/22/16	116	115	1	0
Sørfolda (Sorf)		67°35 N 14°50 E	4/20/16	0	0	0	7
Salten / Skjerstadfj.	Saltenfjord (SALT)	67°16 N 14°38 E	2/15/16	72	190	102	24
	Skjerstadfj. (SKJ)	67°15 N 14°50 E	7/12/16	109			2
Lurefjord (Lure)		60°41 N 05°09 E	6/22/16	188	189	5	22
Raune / Korsfj.	Raunefjord	60°17 N 05°08 E	6/4/16	43	88	0	0
	Korsfjord	60°11 N 05°12 E	6/6/16	45			

Arctic locations are presented first, starting with the northernmost; the Atlantic locations are listed from North to South. Number of individuals analysed is given (“N ind. analysed”), with the precision for the 3 different analyses: “PL” = prosome length measurements; “Redness – Ant / Gen” = examination of redness of antennules / genital somite; “Legs / Gnath” = examination of morphology of the 5th pair of legs and mandibular gnathobase.

Table 2: Pearson's *r* calculation for testing the correlation between *Calanus* body size (prosome length) and latitude.

Species	*C. finmarchicus*			*C. glacialis*		
Stage	CIV	CV	CVI-F	CIV	CV	CVI-F
n	21	161	92	201	269	151
Pearson's *r*	0.8**	0.39**	0.19	0.74**	0.65**	0.84**

Significance levels (p-value) are indicated by: "*": $p < 0.05$, and "**": $p < 0.01$.

Table 3: Antennules red pigmentation in copepodite stages CIV, CV and adult females (CVI F) and males (CVI M) of *Calanus finmarchicus* (*C. fin*) and *C. glacialis* (*C. gla*) from different geographical locations.

a. Nansen Basin		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. fin	CV	100%	0	0	0	23
	CVI F	88%	12%	0	0	24
	Total	94%	6%	0	0	47
C. gla	CVI F	0	4%	85%	11%	47

b. Svalbard		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. fin	CIV	100%	0	0	0	1
	CV	40%	20%	20%	20%	10
	CVI F	59%	35%	6%	0	49
	CVI M	100%	0	0	0	1
	Total	58%	31%	8%	3%	61
C. gla	CIV	0	1%	9%	90%	98
	CV	2%	4%	21%	73%	56
	CVI F	8%	25%	33.5%	33.5%	12
	Total	1%	4%	15%	80%	166

c. White Sea		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. gla	CIV	93%	7%	0	0	100
	CV	0	57%	36%	7%	14
	CVI F	0	0	100%	0	1
	Total	81%	13%	5%	1%	115

d. Salten / Skjerstadfj.		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. fin	CIV	67%	0	0	33%	3
	CV	100%	0	0	0	33
	CVI F	90%	10%	0	0	20
	CVI M	100%	0	0	0	4
	Total	95%	3%	0	2%	60
C. gla	CIV	0	0	0	1	1
	CV	20%	20%	38%	22%	40
	CVI F	4%	16%	40%	40%	84
	CVI M	100%	0	0	0	5
	Total	12%	16%	38%	34%	130

e. Lurefjord		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. fin	CIV	0	0	100%		3
	CV	95%	0	5%		22
	Total	84%	0	16%		25
C. gla	CIV	0	0	100%		1
	CV	25%	0	75%		158
	CVI F	100%	0	0		5
	Total	27%	0	73%		164

f. Raune / Korsfj.		Antennules Redness				
Species	Stage	< 10 %	10-50 %	50-90 %	> 90 %	Total
C. fin	CIV	14%	8%	64%	14%	14
	CV	63%	18%	18%	1%	74
	Total	56%	16%	25%	3%	88

Table 4: Redness of *Calanus finmarchicus* and *C. glacialis* females genital somite. Any red pigmentation observed on one or both spermathecae (= genital field) was reported as "Red". No redness at all was reported as "Pale".

Species	*C. finmarchicus*			*C. glacialis*		
Genital somite	**Red**	**Pale**	Total	**Red**	**Pale**	Total
Svalbard	0	100%	49	73%	27%	11
White Sea	0	0	0	100%	0	1
Salten / Skjerstadfj.	0	100%	20	90%	10%	82
Lurefjord	0	0	0	80%	20%	5
Total	0	100%	69	88%	12%	99

Figures

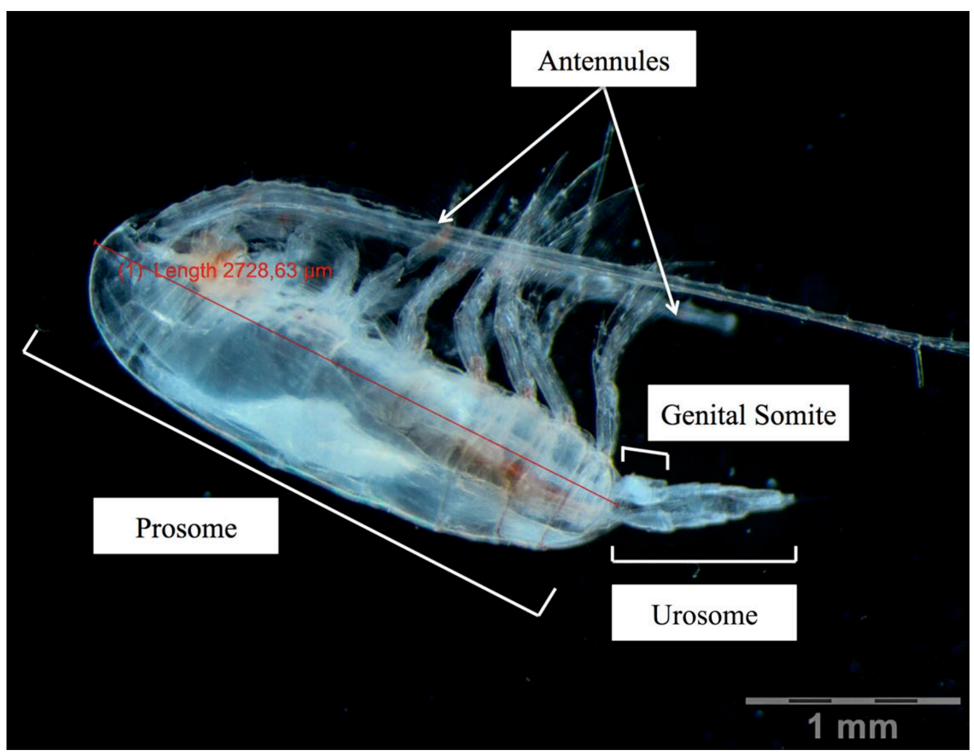


Figure 1: ***Calanus finmarchicus*, structure of the body**

Figure 2: Categories of red pigmentation of antennules in *Calanus*. Two photos are shown as examples for each of the four categories defined as follows: less than 10% of red pigmentation; between 10 and 50% pigmentation; between 50 and 90% of pigmentation; and more than 90% pigmentation.

Figure 3: Pigmentation of genital somite in *Calanus*. Pigmentation is defined as red (left photo) or pale (right photo).

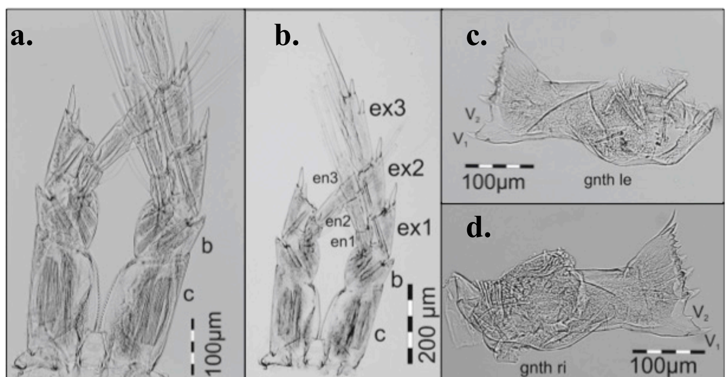


Figure 4: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus finmarchicus* (genetically confirmed) exhibiting the species-specific traits as described in literature. Specimen collected from Van Mijenfjord (ID: VM41). **a** & **b.** Anterior view of the fifth thoracic leg (P5) with denticulated lamellae on the medial margin of the coxa, showing typical "straight form". Abbreviations used: coxa (c); basis (b); - exopods 1 to 3 (ex1-ex3); - endopods 1 to 3 (en1-en3). **c** & **d**. Anterior view of mandible gnathobases, with a typical small second ventral tooth on the cutting edge. Abbreviations: left mandible gnathobase (gnth le); right mandible gnathobase (gnth ri); first ventral tooth (V1); second ventral tooth (V2).

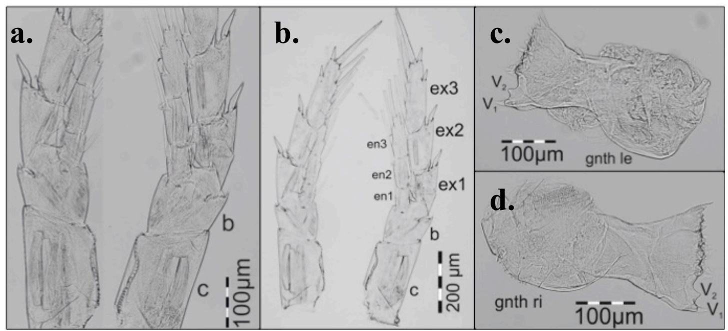


Figure 5: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus glacialis* (genetically confirmed) exhibiting the species-specific traits as described in literature. Specimen collected from Skjerstadjord (ID: SKJ25). **a** & **b.** Anterior view of the 5th thoracic leg (P5) showing denticulated lamellae on the medial margin of the coxa in a typical concave form, with well-expressed curvature. **c** & **d**. Anterior view of mandible gnathobases with the cutting edge with a typical large second ventral tooth on a wide basis. (See legend Fig. 4 for abbreviations meaning).

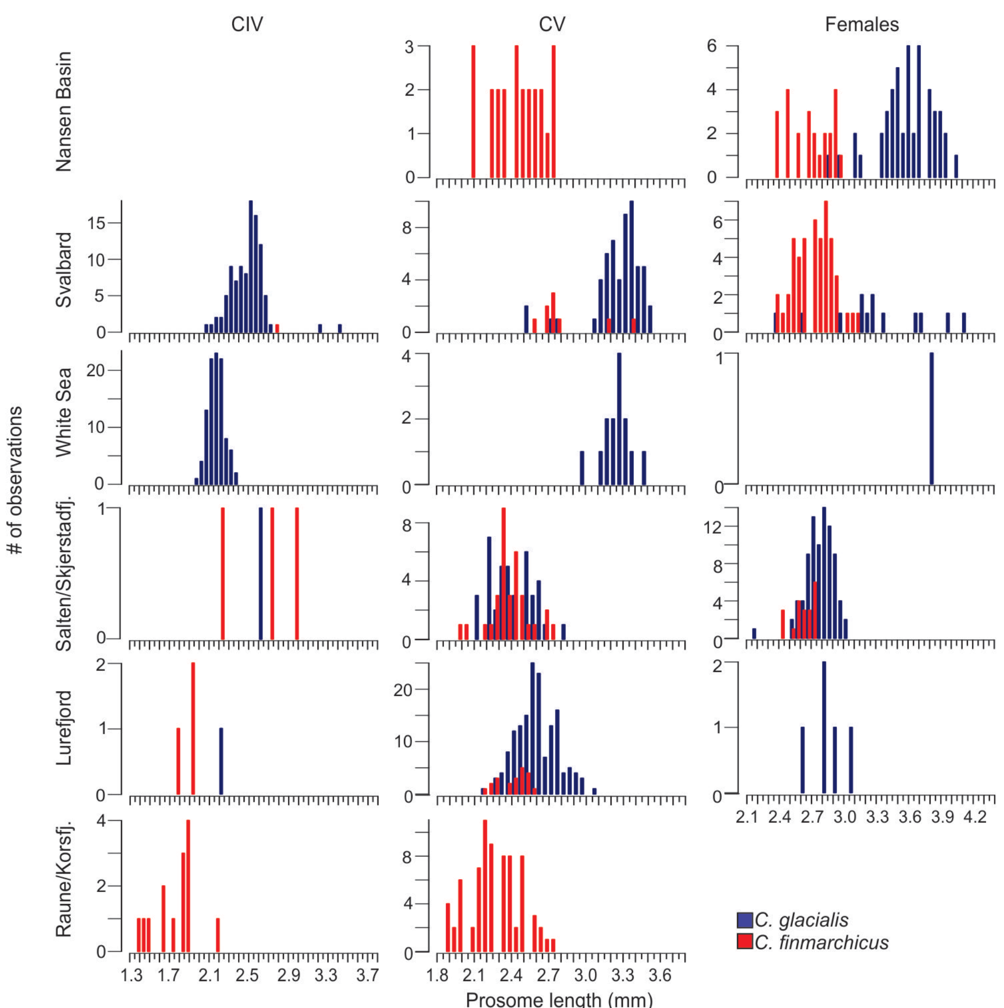


Figure 6: Stage-specific length frequency distributions of prosome length (mm) for copepodites CIV, CV and adult females of *Calanus glacialis* and *C. finmarchicus* in different regions. In total, 895 individuals were measured, from 9 locations, pooled into 6 distant regions, in the North Atlantic and Arctic Oceans. Only *C. glacialis* (blue) occurred in the White Sea, and only *C. finmarchicus* (red) occurred in Raunefjord / Korsfjord area.

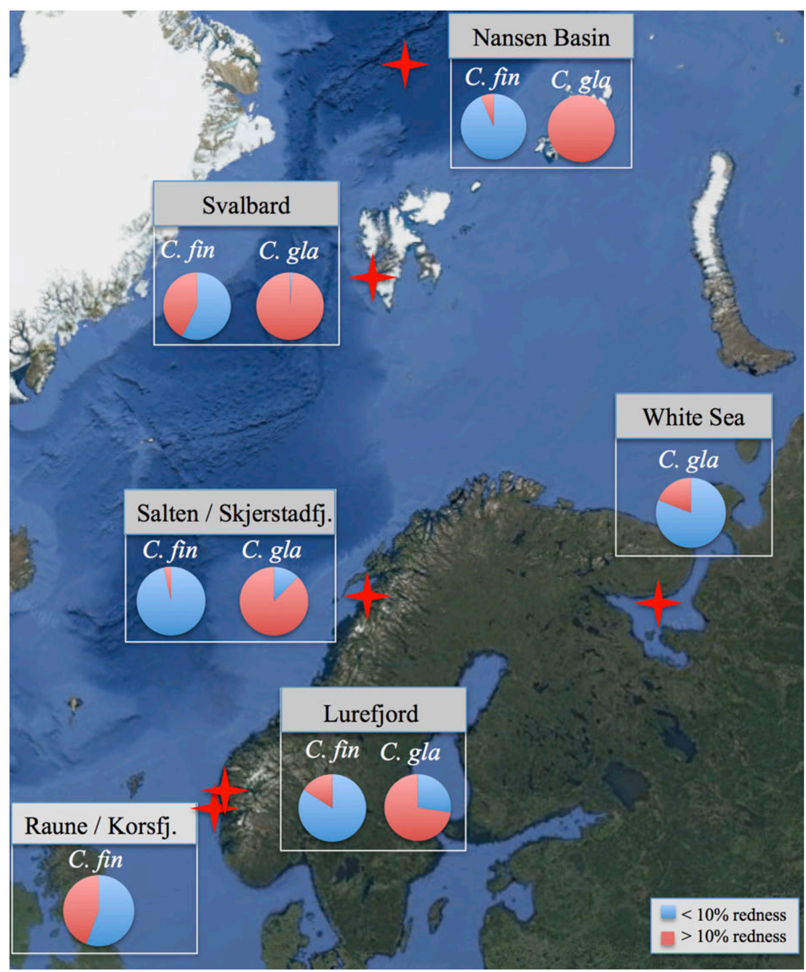


Figure 7: Red pigmentation on *Calanus finmarchicus* (*C. fin*) and *C. glacialis* (*C. gla*) antennules in different regions. Species-related redness from 4 regions where both species co-occur: the Nansen Basin, Svalbard, Saltenfjord / Skjerstadfjord and Lurefjord; and from the White Sea where only *C. glacialis* occurs, and Raunefjord / Korsfjord where only *C. finmarchicus* occurs. Blue colour of the pie charts indicates proportion of individuals for which less than 10% of the surface of their antennules was red; red colour indicates proportion of individuals for which more than 10% of red pigmentation was noticed.

Fig. 8: Antennules redness frequency distribution per *Calanus* species. This violin graph was realised under RStudio v.1.0.143 with the package ggplot2 (Wickham, 2009). The graph shows the distribution of each species individuals on the following 3 ranks scale of redness: 1 = less than 10% of redness; 2 = 10-50% redness; 3 = more than 50% redness.

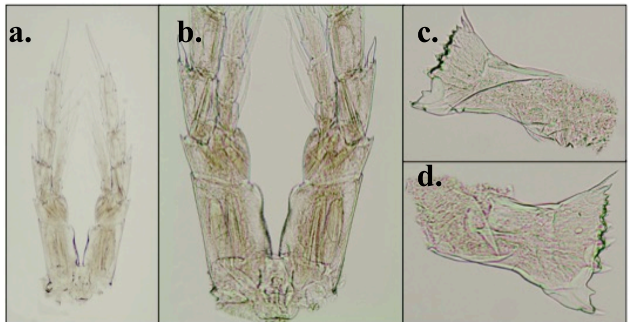


Figure 9: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus finmarchicus* exhibiting traits theoretically assigned to *C. glacialis*. The specimen from Saltenfjord (ID: SALT27) exhibits concave denticulated lamellae with a well-expressed curvature, and a wide basis of the second ventral tooth, typical of *C. glacialis* according to literature.

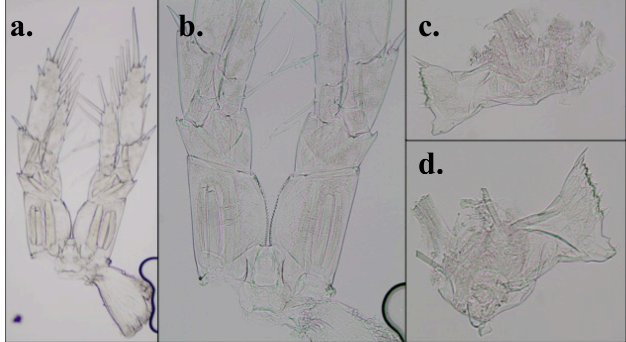


Figure 10: Morphology of the fifth thoracic leg and the gnathobase of an adult female *Calanus glacialis* exhibiting traits theoretically assigned to *C. finmarchicus*. The specimen from Saltenfjord (ID: SALT14) exhibits straight shaped denticulated lamellae, and a small second ventral tooth on the cutting edge of the coxa, typical of *C. finmarchicus* according to literature.

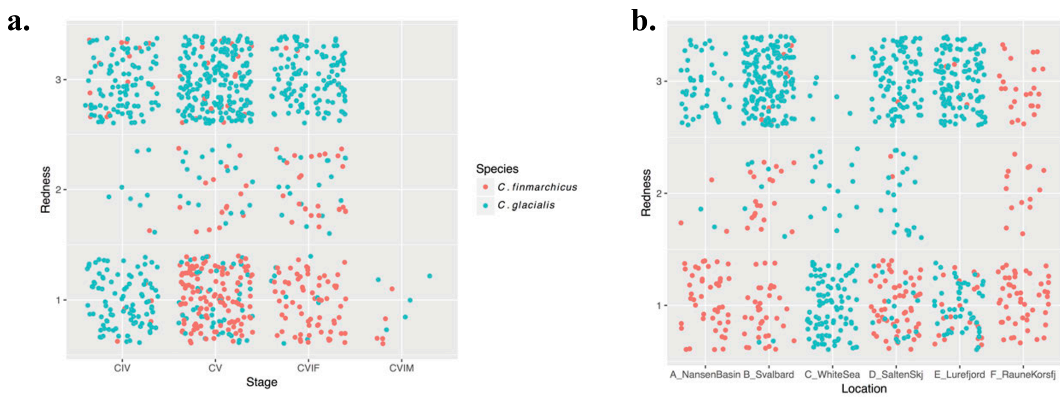
Supplementary Materials

Supplementary 1: Kruskal-Wallis comparison of red pigmentation between species, among developmental stages, and among sampling sites for *Calanus* species.

Kruskal-Wallis	Chi-squared	df	*p*-value
Redness - Species	230.22	1	< 0.01
Redness - Stage	15.502	3	< 0.01
Redness - Stage *C. finmarchicus*	40.121	3	< 0.01
Redness - Stage *C. glacialis*	75.683	3	< 0.01
Redness - Location *C. finmarchicus*	47.522	4	< 0.01
Redness - Location *C. glacialis*	291.03	4	< 0.01

(Results of Kruskal-Wallis H test). "df" stands for the degree of freedom. Calculations were made on RStudio v.1.0.143.

Supplementary 2: Antennules redness frequency distribution per sampled location (a), and per developemental stage (b) for *Calanus finmarchicus* and *C. glacialis.* In panel **a**, 903 individuals are represented grouped per sampled location, in regards to their antennules degree of redness. All stages are mixed together for each location. The sample from "Raune_Korsfj." location contained only *C. finmarchicus* individuals; while "White_Sea" location contained only *C. glacialis* individuals. In panel **b,** 903 individuals are represented grouped by stages CIV, CV, adult females CVI F and adult males CVI M, in regard to degree of redness of their antennules. Degree of redness is ranked from 1 to 3, with 1 = < 10% redness; 2 = 10-50% redness; 3 = > 50% redness.



Supplementary 3:

Stage	Indiv.	P5 legs denticulated margin curvature	Mandible cutting edge, V2	K.K.	S.K.	K.K. VS S.K.	Genetic ID
CV	Lure101	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	Lure102	straight	narrow/small	*C. fin*	*C. fin*	match	*C. gla*
CV	Lure103	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CV	Lure104	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	Lure105	slightly concave	wide/large	*C. fin*	*C. gla*		*C. gla*
CV	Lure107	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	Lure108	concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	Lure110	slightly concave	narrow/small	*C. fin*	*C. fin*	match	*C. gla*
CV	Lure111	slightly concave	wide/large	*C. fin*	*C. gla*		*C. gla*
CV	Lure112	concave	wide/small	*C. gla*	*C. gla*	match	*C. gla*
CV	Lure113	concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	Lure139	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure145	slightly concave	absent/broken	*C. fin*	*C. gla*		*C. fin*
CV	Lure152	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure158	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure183	slightly concave	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure189	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure26	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure27	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure79	slightly concave	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	Lure99	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CV	SALT1	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT10	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CV	SALT11	concave	narrow/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT13	straight	narrow/small	*C. fin*	*C. fin*	match	*C. gla*
CV	SALT14	straight	narrow/small	*C. fin*	*C. fin*	match	*C. gla*
CV	SALT15	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT16	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT17	concave	wide/small	*C. gla*	*C. gla*	match	*C. gla*
CV	SALT18	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT19	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT2	straight	wide/small	*C. fin*	*C. gla*		*C. fin*
CV	SALT20	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT21	slightly concave	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT22	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CV	SALT23	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT24	concave	wide/large	*C. fin*	*C. gla*		*C. gla*
CV	SALT26	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT28	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	SALT31	concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT37	concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT40	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT45	slightly concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CV	SALT6	slightly concave	wide/large	*C. fin*	*C. gla*		*C. gla*
CV	VM10	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CV	VM12	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CV	VM21	slightly concave	narrow/small	*C. gla*	*C. fin*		*C. gla*
CV	VM33	concave	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CV	VM51	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CVIF	Lure109	concave	wide/small	*C. gla*	*C. gla*	match	*C. gla*
CVIF	SALT27	concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	SKJ25	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CVIF	SKJ85	concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	Sorf105	concave	wide/large	*C. gla*	*C. gla*	match	*C. gla*
CVIF	Sorf117	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	Sorf119	concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	Sorf141	concave	wide/small	*C. fin*	*C. gla*		*C. gla*
CVIF	Sorf164	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*
CVIF	Sorf181	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*
CVIF	Sorf191	concave	wide/large	*C. fin*	*C. gla*		*C. gla*
CVIF	VM13	concave	wide/small	*C. gla*	*C. gla*	match	*C. gla*
CVIF	VM15	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	VM16	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*
CVIF	VM29	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	VM34	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*
CVIF	VM4	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*
CVIF	VM41	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CVIF	VM49	straight	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
CVIF	VM5	slightly concave	wide/small	*C. fin*	*C. gla*		*C. fin*
CVIF	VM52	concave	narrow/small	*C. fin*	*C. fin*	match	*C. fin*
	VM76	slightly concave	wide/large	*C. fin*	*C. gla*		*C. fin*

Supplementary 3:

Evaluation of concordance between identification of *Calanus* species by examination of morphological structures and genetics.

Independently from each other, K. Kosobokova (K.K.) and S. Kwaśniewski (S.K.) examined the structure of the 5th pair of legs and of mandible gnathobase of individuals from CV and CVI females for morphological species identification, afterwards verified by genetics ("Genetic ID"). Grey shaded areas represent individuals where both experts agreed on species ID and were confirmed by genetics.

Paper III

Short note – Polar Biology

No evidence of hybridization between the co-occurring copepods *Calanus finmarchicus and C. glacialis*

Choquet M[1,2*], Søreide JE[2], Burckard G[1], Skreslet S[1], Hoarau G[1]

[1]*Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway*
[2]*Department of Arctic Biology, The University Centre in Svalbard, Longyearbyen, Norway*

Abstract

Calanus species and stage composition were analysed using molecular markers (InDels) in two boreal Norwegian fjords where *C. helgolandicus*, *C. finmarchicus*, *C. glacialis* and *C. hyperboreus* co-exist during the main mating season (winter-spring). In these fjords, *Calanus glacialis* were the most abundant, followed by *C. finmarchicus* and *C. hyperboreus* with only a few records of *C. helgolandicus.* Females of *C. glacialis* and *C. hyperboreus* were present in November, while females of *C. finmarchicus* first appeared in January. Males were found for all the three main *Calanus* species in January and February, together with females, suggesting there was an ecological potential for hybridization. However, genetic admixture analyses never detected any signal of hybrids in none of the two fjords suggesting strong isolation mechanism, including species-specific behaviour and/or morphological barriers.

Keywords: *Calanus*, reproduction, Skjerstadfjorden, Mistfjorden, InDels, timing, males

Introduction

Copepods of the genus *Calanus* play a key role in the northern hemisphere as grazers and prey in the marine food web (Falk-Petersen et al., 2009). In the North Atlantic and in the Barents Sea three species prevail: the boreal-Arctic *C. finmarchicus* with main preference for the North Atlantic, the circumpolar Arctic shelf species *C. glacialis*, and the largest of the three the Arctic *C. hyperboreus* with its core distribution in the Greenland Sea (Conover and Huntley, 1991, Hirche, 1989). In addition, the boreal *C. helgolandicus* can be present in low abundances (Choquet et al. subm.). These four species are similar morphologically and have comparable life histories, but differ greatly in their abundances depending on the environment. The *Calanus* species are used as indicator species for specific water masses and temperature regimes and are thus popular study objects to investigate potential impacts of global warming on marine ecosystems (Beaugrand et al., 2003, Beaugrand et al., 2002, Slagstad et al., 2011).

Despite being among the most studied organisms within the zooplankton, *Calanus* identification to species level remains a challenge, and particularly to discriminate between *C. glacialis* and *C. finmarchicus* in areas of sympatry (Choquet et al. subm., Gabrielsen et al., 2012, Lindeque et al., 2006, Parent et al., 2011). Therefore, different molecular tools have been developed in order to facilitate the identification of these important species without bias due to morphological plasticity (Lindeque et al., 1999, Provan et al., 2009, Smolina et al., 2014). Currently, the easiest and fastest approach for reliable species identification within *Calanus* genus consists in the use of 6-12 markers type insertion/deletion (InDel) (Smolina et al., 2014). These markers are nuclear and co-dominant and thus allow detection of potential hybrids. Their performance has been tested over thousands of individuals across the North Atlantic and Arctic, and their validity has been confirmed by comparison with traditional 16S mtDNA sequencing for species identification (see Choquet et al., subm., Nielsen et al., 2014, Smolina et al., 2014).

Recent studies based on molecular markers type microsatellites suggested that *C. finmarchicus* and *C. glacialis* are able to interbreed to produce fertile hybrids at high rates in the northwest Atlantic and the Canadian Arctic (Parent et al., 2015, Parent et al., 2012). In West and East Greenland, the same microsatellites were combined with InDels to search for hybridization between *Calanus finmarchicus* and *C. glacialis* (Nielsen et al., 2014). No hybrids were detected and simulations showed that the microsatellites alone had less power to fully discriminate between the introgressed individuals and parental species, compared to the InDels (Nielsen et al., 2014).

In the present study, we combined classical ecology with use of these new molecular InDels markers to test for hybridization between *C. glacialis* and *C. finmarchicus* in two Norwegian fjords where they live in sympatry. During the main *Calanus* mating season from November to March we regularly sampled the *Calanus* community to investigate the ecological potential for hybridization, simultaneously with molecular analyses to detect putative hybridization.

Materials and Methods

Sample collection

Zooplankton samples were collected in Skjerstadfjord (0 – 500 m) and Mistenfjord (0 – 285 m), in Northern Norway (Fig. 1) using a Juday net with mesh size of 200 μm and net opening 0,096 m^2. Measurements of temperature, salinity, oxygen and fluorescence were carried out by a CTD at each sampling. We assumed 100% filtration efficiency of the net. Sampling was done with a periodicity of approximately one month between November 2016 and March 2017 in Skjerstadfjord (16/11/2016, 21/12/2016, 30/01/2017, 28/02/2017, 30/03/2017) and Mistenfjord (17/11/2016, 08/12/2016, 24/01/2017, 23/02/2017, 29/03/2017). Five replicates were collected for each date and location. Samples were preserved in 70-80% undenaturated ethanol, with subsequent change of ethanol after the 24 first hours.

Stage composition

Samples of the 3 first replicates of each month were divided using a Folsom plankton splitter in subsamples containing about 50 *Calanus* individuals. 48 of these individuals were identified to developmental stage and sex under a stereomicroscope (Leica 10X /23, X4). They comprised almost exclusively the overwintering stages CIV, CV and adults (CVI). The very few CIII present were discarded and we divided the remaining into 4 groups: CIV, CV, CVI males and CVI females in 4 Petri plates (one for each developmental stage and sex) containing nuclease-free water in order to remove the ethanol, and gave them a unique ID to proceed to the DNA extraction.

Molecular species identification

Insertion – deletion polymorphism (InDel) was used to identify each individual as *Calanus finmarchicus*, *C. glacialis*, *C. hyperboreus* or *C. helgolandicus*. For each replicate, we removed the antennules of the 48 individuals selected beforehand to extract the DNA of each specimen, using the quick and cheap method of HotSHOT DNA extraction (Montero - Pau et al., 2008).

We then amplified a set of 6 nuclear InDel markers (Smolina et al., 2014) in a single multiplexed Polymerase Chain Reaction (PCR) using a Veriti 96-Well Fast Thermal Cycler (Life Technologies), following the protocol described by Smolina *et al*. (2014). Cycling parameters for PCR were: one initial denaturation at 94°C (2 min), followed by 35 cycles of 94°C (10 s), 55°C (10 s), 72°C (10 s) and one final extension cycle at 72°C.

To prepare the genotyping, we diluted the PCR products adding 45μL of nuclease-free water in each well of the PCR plate; we then transferred 2μL of the diluted products in another PCR plate containing 38μL of nuclease-free water in each well. 1 μL of the diluted DNA of each specimen was transferred in a reaction plate with 9μL of a mixture Formamide (8.9 μL) – Liz standard 500 (0.1μL). The plate was placed in a thermocycler (TC-412 - TECHNE) to denature the DNA (5 min – 95°C). Fragments analysis was done following the protocol described by Smolina *et al*. (2014) using a 3500xL Genetic Analyser (Life Technologies). Using GENEMAPPER 2.4.1 (Life Technologies), four distinct patterns of genotypes were distinguished and assigned to the four different species of *Calanus* based on species-specific alleles defined in Smolina *et al*. (2014).

Putative hybridization was investigated using an admixture analysis following Nielsen et al. (2014). We used STRUCTURE (V. 2.3.4) (Pritchard et al., 2000) to analyse the admixture of nuclear genotypes (parameters: ancestry model = admixture; frequency model = correlated; burn-in = 2 000 000; MCMC length = 1 000 000 after burn-in). The software uses a Bayesian algorithm to identify K (K=2 for *C. finmarchicus* and *C. glacialis*) clusters of genetically homogenous individuals. Based on their multi-locus genotypes, each individual is then characterised by admixture coefficient, defined as the probability of belonging to the *C. finmarchicus* or *C. glacialis* cluster.

Results

Physical and biological environment

Skjerstadfjorden is a deep (500 m) fjord with a very narrow and shallow sill (23 m) that severely restricts the water exchange with Saltfjorden and the shelf seas outside (Fig.1). Mistfjorden is also a sill fjord but the sill is deeper (34 m), the maximum depth shallower (285 m) and the fjord smaller in area compared to Skjerstadfjorden (Fig.1). No indication of an advective event was detectable from the hydrographical measurements (Supplementary 1). Both fjords were dominated by local waters with warm (ca. 8°C) and relatively fresh (<32 psu) surface waters in November that gradually cooled and become more saline as winter convection proceeded towards March (Supplementary 1). Below 50 to 150 m the hydrography was relatively stable

over the period sampled with temperatures around 5-6°C and salinities between 34.2-34.3 psu in both fjords. Fluorescence measurements showed typically low winter chlorophyll a values throughout the entire period sampled.

Species composition

The *Calanus* abundances were slightly lower in Skjerstadfjorden (1300-1450 ind. m^{-2}) than in Mistfjorden (1700-2300 ind. m^{-2}), but all four species were identified in both fjords (Fig.2). The occurrence of *C. helgolandicus*, however, was very low. In Skjerstadfjorden, *C. glacialis*, *C. finmarchicus* and *C. hyperboreus* were equally present from November to January. In February, the numbers of *C. finmarchicus* started to decline and in March its abundance was similarly low as that of *C. helgolandicus*. For *C. glacialis* and *C. hyperboreus* the population numbers were constant from November to February, but a marked decline was also seen for these two in March.

In Mistfjorden, *C. glacialis* dominated in term of abundance (50-70%), followed by *C. finmarchicus* (~20%), *C. hyperboreus* (5-20%) and *C. helgolandicus* (0-2%) (Fig.2). However, also here a strong decline in population numbers was found in March.

Stage composition

The dominant developmental stage was CV for all species in both fjords from November to January (Fig.2). Females of *C. glacialis* and *C. hyperboreus* were found from November on, while from January for *C. finmarchicus* and *C. helgolandicus*. Females peaked in abundance for all four species in February-March. Males of *C. glacialis*, *C. finmarchicus* and *C. hyperboreus* appeared in January and prevailed until end of February for *C. glacialis* and *C. hyperboreus*, and to March for *C. finmarchicus*. In March, also a few males of *C. helgolandicus* were recorded in Skjerstadfjorden, otherwise not.

Hybridization

The timing of male and female occurrence revealed a potential for hybridization between *C. finmarchicus* and *C. glacialis*. The genetic admixture analysis, however, did not detect any evidence of hybridization in Skjerstadfjorden (Fig.3) nor Mistfjorden (Supplementary 3).

Discussion

The absence of hybrids between the morphologically similar sibling species *C. glacialis* and *C. finmarchicus* in this study and elsewhere in the North Atlantic, West

and East Greenland and the high-Arctic (Choquet et al. subm., Nielsen et al. 2014) strongly suggests that these two species do not hybridize, despite that both species have males and females present at the same time.

The previous study where hybrids of *C. glacialis* and *C. finmarchicus* were reported was based on microsatellites markers (Parent et al., 2015, Parent et al., 2012). Microsatellites may not be the ideal type of marker for species identification and to detect hybridization because of common occurrences of null alleles (Dakin and Avise, 2004), possible homoplasy when comparing two species (Chambers and MacAvoy, 2000), high mutation rate and difficulties to score alleles (Pompanon et al., 2005, Selkoe and Toonen, 2006). In contrast, InDel markers have a low mutation rate, due to single mutation event, thus resulting in a conserved phylogenetic signal (Liu and Cordes, 2004) and alleles easier to genotype with more reproducibility (Väli et al., 2008). Indeed, the newly developed *Calanus* InDels have shown to better discriminate between the two species (Nielsen et al., 2014, Smolina et al., 2014).

Rare or past introgression may not be detectable with a small number of markers (10s) and would require large next-generation sequencing datasets to be detected (Martin and Jiggins, 2017, Rosenzweig et al., 2016). However, the presence of 1st generation hybrids (F1) can be excluded, both in the present data set and in the extensive genotyping carried out on *Calanus* spp. in the North Atlantic (>4400 individuals, Choquet et al. subm.). Furthermore, no nucleocytoplasmic disequilibrium (signature of past hybridization) was found in the 677 individuals genotyped both for nuclear (InDel) and mtDNA (16S) in West and East Greenland (389 in Nielsen et al., 2014) and in the North Atlantic (288 in Choquet et al. subm.).

Parent et al. (2012, 2015) reported high proportions of hybrids and suggested that these hybrids are fertile. If true, this should lead to large-scale introgression and ultimately to the formation of hybrid swarms (e.g. Perry et al., 2001). However, neither was found in the *Calanus* analyse in this or in recent studies (Nielsen et al., 2014, Choquet et al., subm.) so species appear to have remained genetically distinct (Choquet et al subm.).

Several mechanisms are likely to contribute to the species boundaries in *Calanus*. Males and females must be present at the same time to mate. Newly moulted females signal to males by depositing vertical pheromone trails that males search for and follow (Tsuda and Miller, 1998). Hybridization may simply not take place due to very species-specific pheromones or too large of a distance between males and females of different species due to different depth preferences and/or timing in seasonal ascent. Morphological studies have documented on the similarity of sexually modified appendages and body segments of *Calanus* species (Brodskii, 1967, Frost, 1974).

However, strong species-specific differences in the ventral integumental organs of the female urosome have also been reported and suggested to have a role in protecting the sibling species from hybridization (Fleminger and Hulsemann, 1977). Our knowledge on *Calanus* pheromones and the actual mating with copulatory clasp and spermatophore transfer are basically non-existent for *Calanus* and more studies are critically needed.

In Skjerstadfjorden and Mistfjorden, males of *C. hyperboreus*, *C. glacialis* and *C. finmarchicus* first appeared in January and by the end of February they were already gone, except for males of *C. finmarchicus.* Monthly sampling may have been too coarse to detect species-specific differences in timing of male occurrence since males are present for such a short period of time. This is not the case, however, for females that persist for a much longer time in the population. Here, a distinct difference in female-timing was seen between the cold adapted Arctic species *C. glacialis* and *C. hyperboreus* and the more boreal *C. finmarchicus* and *C. helgolandicus*, with the Arctic species preceding the boreal ones by up to two months (November versus January). Maturation of the female gonads is energy and time demanding and is associated with a strong decline in females lipid reserves (Jonasdottir, 1999, Niehoff, 1998). In Skjerstadfjorden and Mistfjorden, the low fluorescence data throughout the entire sampling period from November to March indicated very poor feeding conditions. Thus, gonad maturation had to be fuelled by internal reserves only, which may explain the strong decline in *Calanus* abundance, and especially for *C. finmarchicus* in both fjords from March to February.

Post-zygotic isolation mechanisms could also play a role to prevent hybridization. All studies (including the present paper) have focused on older stages and adults but the early stages are usually ignored due to technical challenges associated with small individuals. Together with whole genome scans, genotyping early stages of *Calanus* in the fjords would be key to further investigate *Calanus* species boundaries and should therefore be prioritised in future hybridization studies together with studies on the *Calanus* mating process.

Acknowledgements

We would like to thank Morten Krogstad for the sampling. We are very grateful to Maja Hatlebakk for processing the CTD data, and to Malin Daase for making the map. This project was supported by the Norwegian Research Council (216578 and 227139) and Nord University.

References

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S. & Reid, P.C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature,* 426**:** 661-664.

Beaugrand, G., Ibañez, F., Lindley, J.A. & Reid, P.C. (2002). Diversity of calanoid copepods in the North Atlantic and adjacent seas: species associations and biogeography. *Marine Ecology Progress Series,* 232**:** 179-195.

Brodskii, K. (1967). Types of female genitalia and heterogeneity of genus *Calanus* (Copepoda). *Doklady Akademii Nauk SSSR,* 176**:** 222-&.

Chambers, G.K. & Macavoy, E.S. (2000). Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology,* 126**:** 455-476.

Conover, R.J. & Huntley, M. (1991). Copepods in ice-covered seas. Distribution, adaptations to seasonally limited food, metabolism, growth patterns and life cycle strategies in polar seas. *Journal of Marine Systems,* 2**:** 1-41.

Dakin, E. & Avise, J. (2004). Microsatellite null alleles in parentage analysis. *Heredity,* 93**:** 504-509.

Falk-Petersen, S., Mayzaud, P., Kattner, G. & Sargent, J. (2009). Lipids and life strategy of Arctic *Calanus*. *Marine Biology Research,* 5**:** 18-39. DOI 10.1080/17451000802512267.

Fleminger, A. & Hulsemann, K. (1977). Geographical range and taxonomic divergence in North Atlantic *Calanus* (*C. helgolandicus*, *C. finmarchicus* and *C. glacialis*). *Marine Biology,* 40**:** 233-248.

Frost, B. (1974). *Calanus marshallae*, a new species of calanoid copepod closely allied to the sibling species *C. finmarchicus* and *C. glacialis*. *Marine Biology,* 26**:** 77-99.

Gabrielsen, T.M., Merkel, B., Søreide, J.E., Johansson-Karlsson, E., Bailey, A., Vogedes, D., Nygård, H., Varpe, Ø. & Berge, J. (2012). Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. *Polar Biol,* 35**:** 1621-1628. DOI 10.1007/s00300-012-1202-7.

Hirche, H.J. (1989). Spatial distribution of digestive enzyme activities of *Calanus finmarchicus* and *C. hyperboreus* in Fram Strait/Greenland Sea *Journal of Plankton Research,* 11**:** 431-443.

Jonasdottir, S.H. (1999). Lipid content of *Calanus finmarchicus* during overwintering in the Faroe-Shetland Channel. *Fisheries Oceanography***:** 61-72.

Lindeque, P.K., Hay, S.J., Heath, M.R., Ingvarsdottir, A., Rasmussen, J., Smerdon, G.R. & Waniek, J.J. (2006). Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. *Journal of Plankton Research,* 28**:** 221-238. DOI 10.1093/plankt/fbi115.

Lindeque, P., Harris, R., Jones, M. & Smerdon, G. (1999). Simple molecular method to distinguish the identity of *Calanus* species (Copepoda: Calanoida) at any developmental stage. *Marine Biology,* 133**:** 91-96.

Liu, Z.J. & Cordes, J. (2004). DNA marker technologies and their applications in aquaculture genetics. *Aquaculture,* 238**:** 1-37.

Martin, S.H. & Jiggins, C.D. (2017). Interpreting the genomic landscape of introgression. *Current Opinion in Genetics & Development,* 47**:** 69-74.

Montero ‐ Pau, J., Gómez, A. & Muñoz, J. (2008). Application of an inexpensive and high ‐ throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods,* 6**:** 218-222.

Niehoff, B. (1998). The gonad morphology and maturation in Arctic *Calanus* species.

Nielsen, T., Kjellerup, S., Smolina, I., Hoarau, G. & Lindeque, P. (2014). Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phenological differences? *Marine Biology,* 161**:** 1299-1306. DOI 10.1007/s00227-014-2419-5.

Parent, G.J., Plourde, S., Joly, P. & Turgeon, J. (2015). Phenology and fitness of *Calanus glacialis*, *C. finmarchicus* (Copepoda), and their hybrids in the St. Lawrence Estuary. *Marine Ecology Progress Series,* 524**:** 1-9. DOI 10.3354/meps11240.

Parent, G.J., Plourde, S. & Turgeon, J. (2012). Natural hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and Northwest Atlantic. *Limnology and Oceanography,* 57**:** 1057-1066. DOI 10.4319/lo.2012.57.4.1057.

Parent, G.J., Plourde, S. & Turgeon, J. (2011). Overlapping size ranges of *Calanus* spp. off the Canadian Arctic and Atlantic Coasts: impact on species' abundances. *Journal of Plankton Research,* 33**:** 1654-1665. DOI 10.1093/plankt/fbr072.

Perry, W.L., Feder, J.L., Dwyer, G. & Lodge, D.M. (2001). Hybrid zone dynamics and species replacement between Orconectes crayfishes in a northern Wisconsin lake. *Evolution,* 55**:** 1153-1166.

Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics,* 6**:** 847-846.

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

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A. & Savidge, G. (2009). High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus. Proceedings Biological sciences / The Royal Society,* 276**:** 301-7. DOI 10.1098/rspb.2008.1062.

Rosenzweig, B.K., Pease, J.B., Besansky, N.J. & Hahn, M.W. (2016). Powerful methods for detecting introgressed regions from population genomic data. *Molecular ecology,* 25**:** 2387-2397.

Selkoe, K.A. & Toonen, R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters,* 9**:** 615-629.

Slagstad, D., Ellingsen, I.H. & Wassmann, P. (2011). Evaluating primary and secondary production in an Arctic Ocean void of summer sea ice: An experimental simulation approach. *Progress in Oceanography,* 90**:** 117-131. DOI 10.1016/j.pocean.2011.02.009.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14**:** 1072-9. DOI 10.1111/1755-0998.12241.

Tsuda, A. & Miller, C.B. (1998). Mate-finding behaviour in *Calanus marshallae* Frost. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences,* 353**:** 713-720. DOI 10.1098/rstb.1998.0237.

Väli, Ü., Brandström, M., Johansson, M. & Ellegren, H. (2008). Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC genetics,* 9**:** 8.

Figures

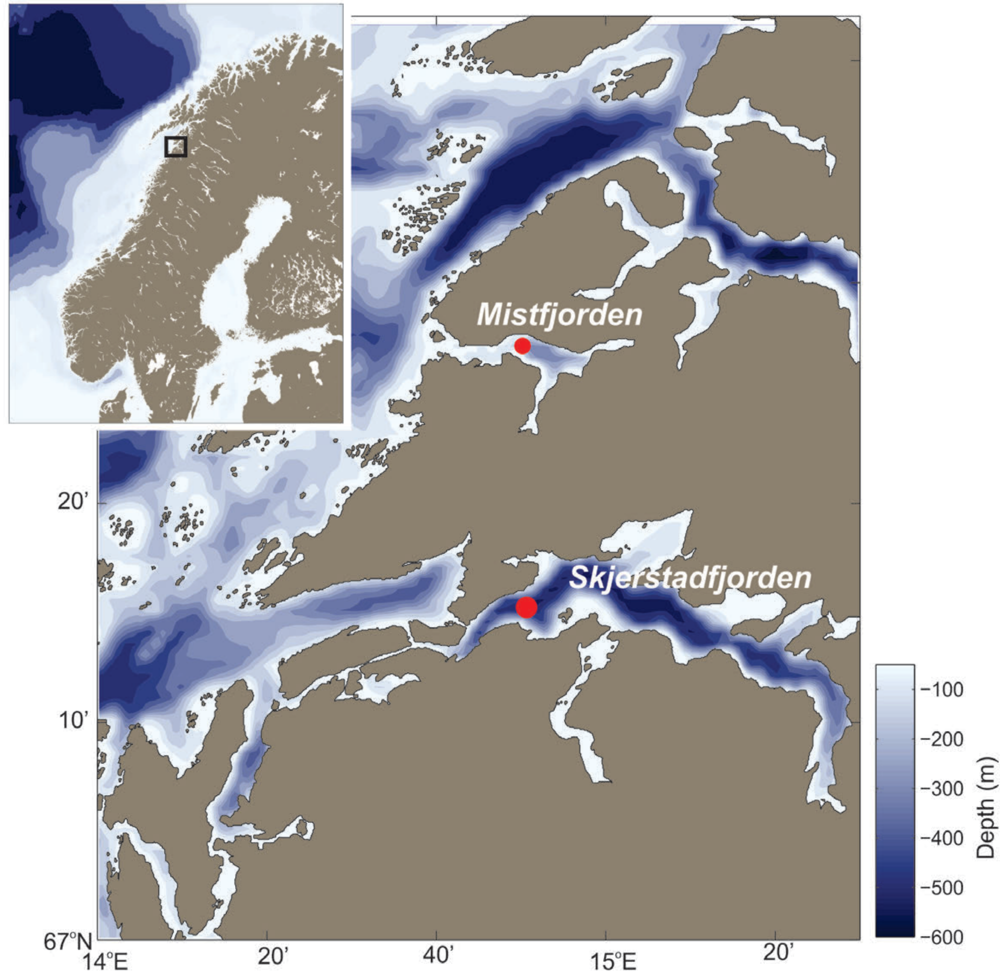


Fig. 1. Map of the sampling locations in Skjerstadfjorden and Mistfjorden in Northern Norway.

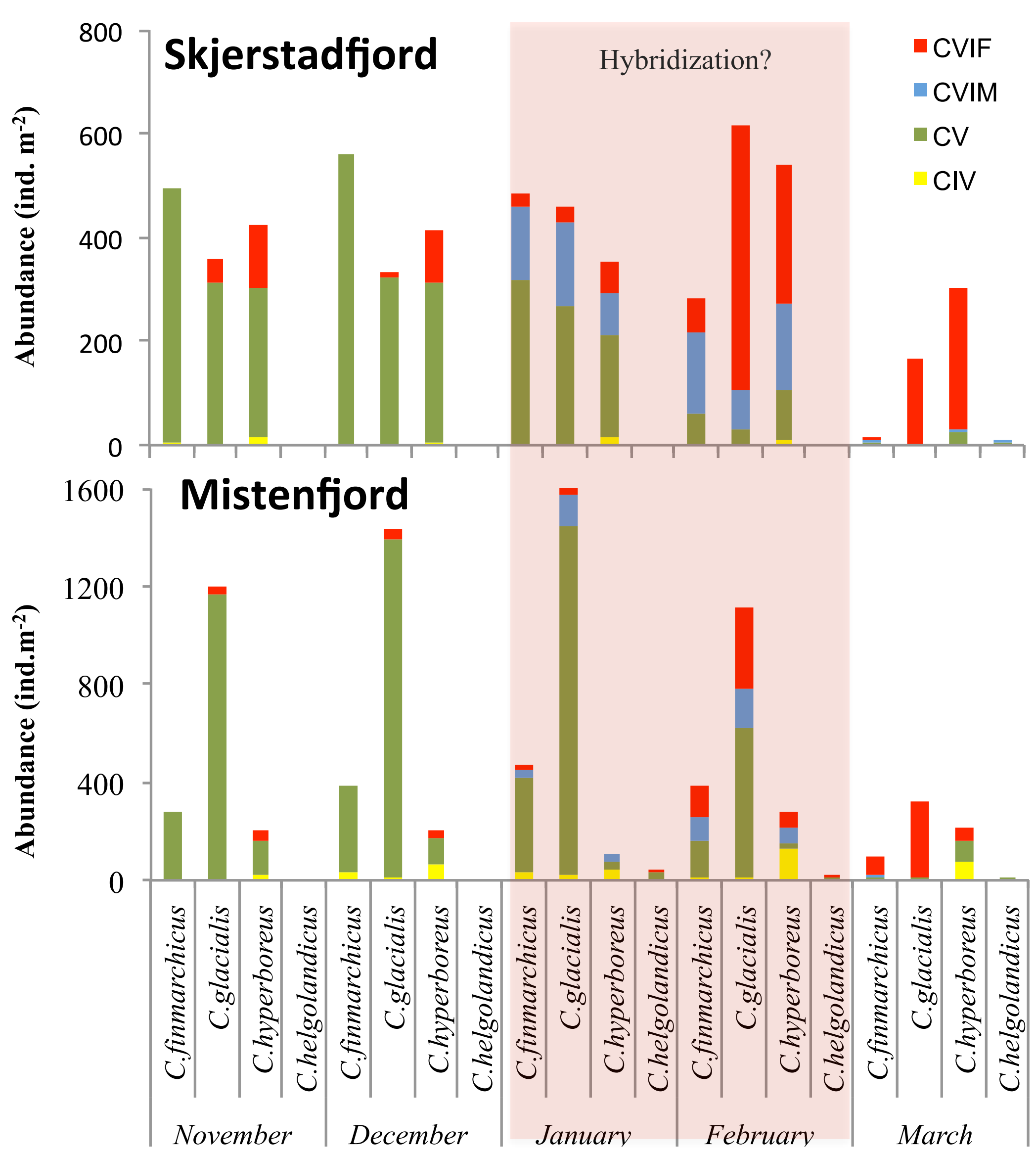


Fig. 2. *Calanus* abundances in Skjerstadfjorden and Mistfjorden from November 2016 to March 2017. Timing of potential hybridization in January and February when both females and males are present is illustrated (red box).

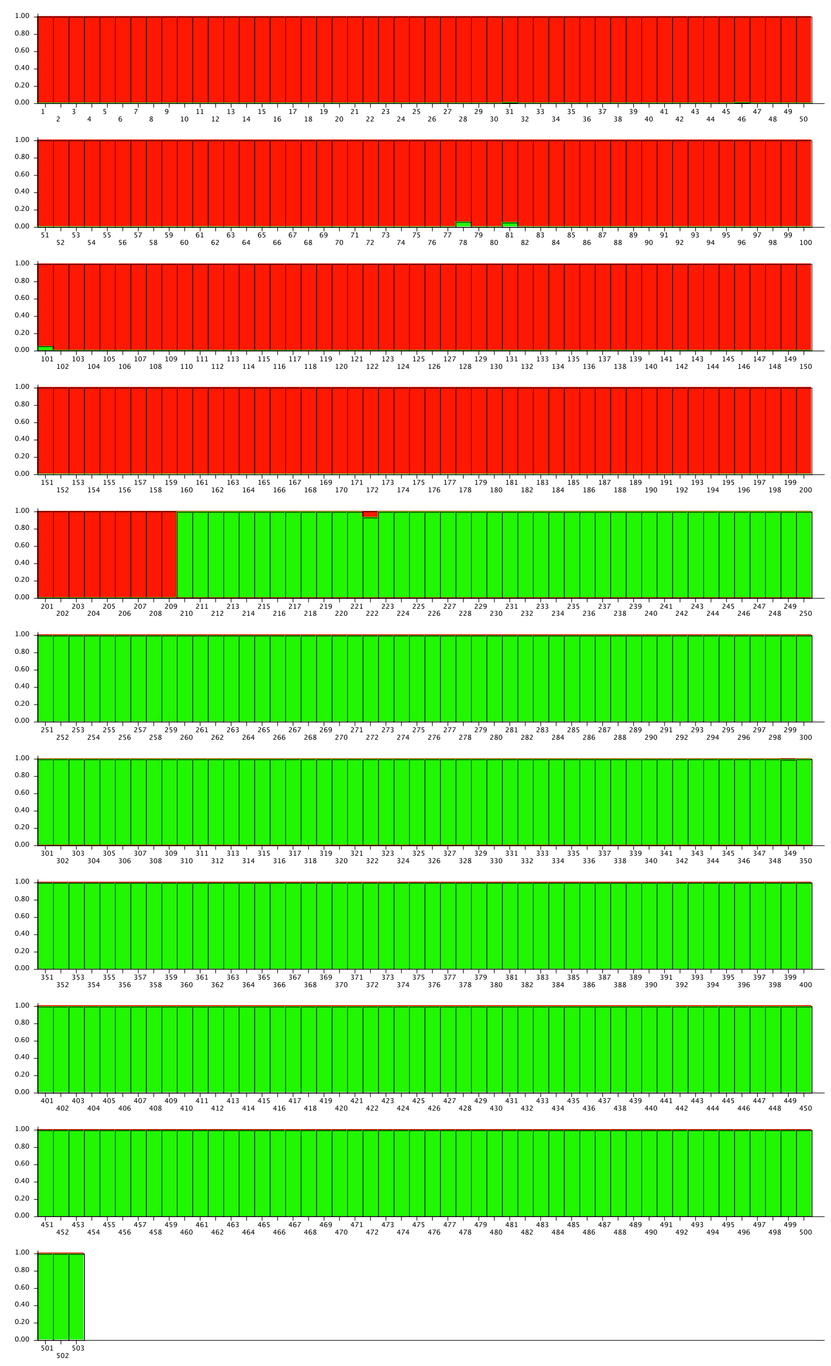
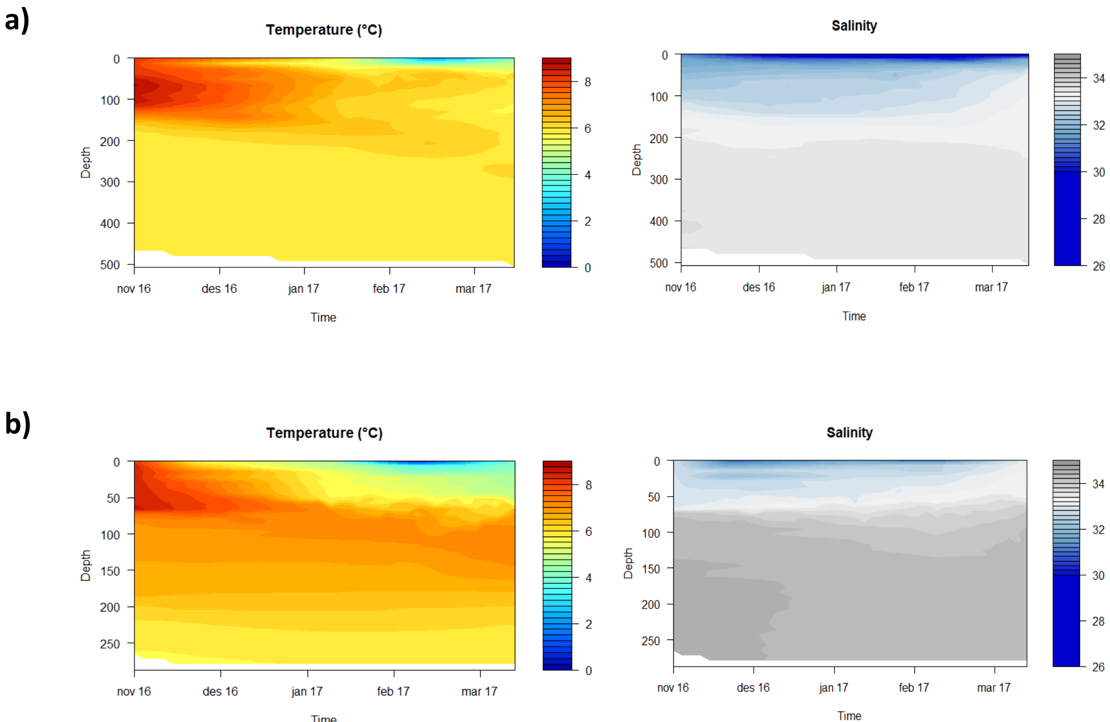
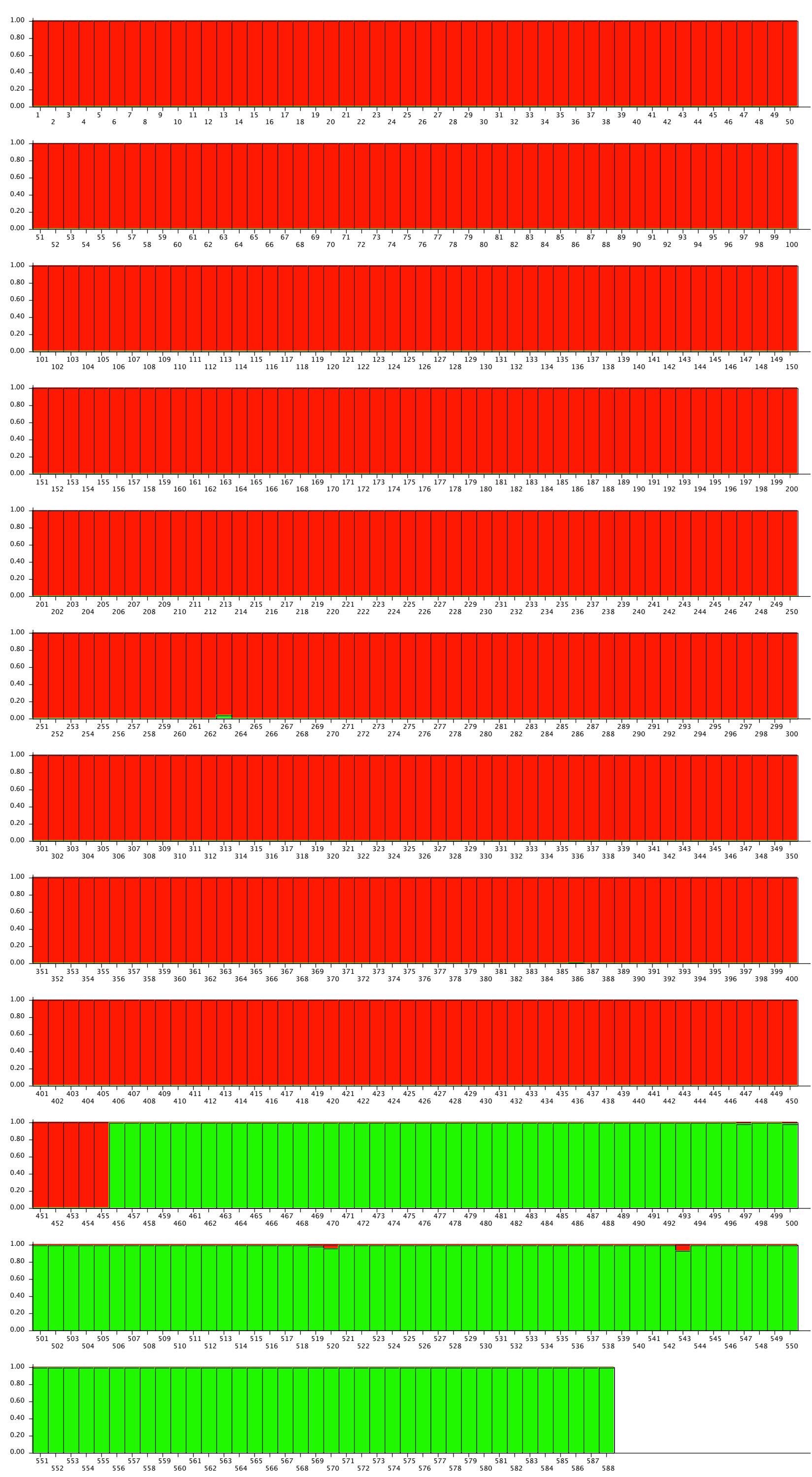


Fig. 3. Genotype admixture analysis based on nuclear InDels markers shows no hybrids between *Calanus finmarchicus* (red) and *C. glacialis* (green) in Skjerstadfjorden.

Supplementary Materials



Supp. 1. Temperature (°C) and salinity (psu) development from November 2016 to March 2017 in a) Skjerstadfjorden and b) Mistfjorden.



Supp. 2. Genotype admixture analysis based on nuclear InDels markers shows no hybrids between *Calanus finmarchicus* (green) and *C. glacialis* (red) in Mistfjorden.

		Skjerstadfjord				
		C4	C5	Male	Female	Total
November (16/11/2016)	*C.finmarchicus*	1	52	0	0	53
	C.glacialis	0	33	0	4	37
	C.hyperboreus	1	34	0	14	49
	C.helgolandicus	0	0	0	0	0
	Total	2	119	0	18	**139**
December (21/12/2016)	*C.finmarchicus*	0	60	0	0	60
	C.glacialis	0	34	0	1	35
	C.hyperboreus	1	33	0	12	46
	C.helgolandicus	0	0	0	0	0
	Total	1	127	0	13	**141**
January (30/01/2017)	*C.finmarchicus*	0	39	17	3	59
	C.glacialis	0	27	16	3	46
	C.hyperboreus	3	18	7	7	35
	C.helgolandicus	0	0	0	0	0
	Total	3	84	40	13	**140**
February (29/02/2017)	*C.finmarchicus*	0	6	15	6	27
	C.glacialis	0	3	7	49	59
	C.hyperboreus	1	9	16	26	52
	C.helgolandicus	0	0	0	0	0
	Total	1	18	38	81	**138**
March (30/03/2017)	*C.finmarchicus*	0	1	2	2	5
	C.glacialis	0	0	0	48	48
	C.hyperboreus	0	8	1	78	87
	C.helgolandicus	0	2	1	0	3
	Total	0	11	4	128	**143**
May (11/05/2017)	*C.finmarchicus*	0	4	1	0	5
	C.glacialis	6	59	0	4	69
	C.hyperboreus	2	18	0	2	22
	C.helgolandicus	0	0	0	0	0
	Total	8	81	1	6	**96**

		Mistfjord				
		C4	C5	Male	Female	Total
November (17/11/2016)	*C.finmarchicus*	0	23	0	0	23
	C.glacialis	0	102	0	3	105
	C.hyperboreus	2	12	0	3	17
	C.helgolandicus	0	0	0	0	0
	Total	2	137	0	6	**145**
December (08/12/2016)	*C.finmarchicus*	2	24	0	0	26
	C.glacialis	1	92	0	3	96
	C.hyperboreus	4	7	0	2	13
	C.helgolandicus	0	0	0	0	0
	Total	7	123	0	5	**135**
January (24/01/2017)	*C.finmarchicus*	2	24	2	1	29
	C.glacialis	1	88	8	4	101
	C.hyperboreus	3	2	2	0	7
	C.helgolandicus	0	2	0	1	3
	Total	6	116	12	6	**140**
February (23/02/2017)	*C.finmarchicus*	1	11	8	10	30
	C.glacialis	1	48	12	26	87
	C.hyperboreus	10	2	5	5	22
	C.helgolandicus	0	1	0	1	2
	Total	12	62	25	42	**141**
March (29/03/2017)	*C.finmarchicus*	0	3	1	21	25
	C.glacialis	0	2	0	64	66
	C.hyperboreus	17	18	0	12	47
	C.helgolandicus	0	1	0	0	1
	Total	17	24	1	97	**139**

Supp. 3. Species and stage composition of *Calanus* spp. in Skjerstadfjorden and Mistfjorden (Number of individuals genetically identified).

Paper IV

This is an open access publication and was reproduced under the terms of the Creative Commons licence.

MITOCHONDRIAL DNA PART B: RESOURCES, 2017
VOL. 2, NO. 2, 506–507
https://doi.org/10.1080/23802359.2017.1361357

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

OPEN ACCESS Check for updates

The complete mitochondrial genome of the copepod *Calanus glacialis*

Marvin Choquet[a], Homère J. Alves Monteiro[a], Johan Bengtsson-Palme[b] and Galice Hoarau[a]

[a]Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway; [b]Department of Infectious Diseases, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ABSTRACT
Calanus glacialis, a marine planktonic copepod, is a keystone species in the Arctic Ocean. In this study, we shotgun sequenced the total DNA of one *C. glacialis* individual, using the NextSeq® Illumina platform, in order to determine its mitochondrial genome sequence. We successfully assembled and annotated this 20,674 bp long sequence, which included 13 protein-coding genes, 2 rRNA genes and 22 tRNA genes. Common gene-coding regions of 19 other species were used to reconstruct a phylogenetic tree, using mitogenomes of the phylogenetically closest copepods available. The new resource described here constitutes a tool of interest for better understanding the structure and dynamics of *C. glacialis* populations.

ARTICLE HISTORY
Received 4 July 2017
Accepted 26 July 2017

KEYWORDS
Calanus glacialis; copepod; mitochondrion; mitogenome; zooplankton

Downloaded by [81.191.247.64] at 16:33 07 August 2017

The genus *Calanus* consists of 26 distinct marine copepod species (WoRMS Editorial Board, 2017), present in every ocean in the world as part of the zooplankton. Despite their ecological importance, only two mitochondrial genomes have been reported within the *Calanus* genus: *C. sinicus* (Minxiao et al. 2011) and *C. hyperboreus* (Kim et al. 2013). *C. glacialis* is one of the key species of the Arctic Ocean, as the crucial link between primary production and higher trophic levels such as fishes, invertebrates, marine mammals and birds (Falk-Petersen et al. 1990).

In this study, we report the complete sequence of the mitochondrial genome of *C. glacialis*. We selected one *C. glacialis* individual from Sørfolda (Norwegian coast: 67°30 N, 15°28 E), which we identified as such using a set of nuclear molecular markers (Smolina et al. 2014). Total DNA was extracted using the E.Z.N.A.® Insect DNA Kit and was shotgun sequenced on the NextSeq® Illumina platform. Given the amount of DNA recovered from a single individual, everything was used for the library construction. *De novo* assembly of the filtered reads was performed using Ray version 2.3.1 (Boisvert et al. 2010) with a k-mer length of 31. Contigs that matched the mitochondrial genomes of *C. hyperboreus* or *C. sinicus* in a BLAST (Altschul et al. 1997) search (e-value cut-off 10^{-10}) were extracted. To potentially further merge these contigs, they were used as seeds in a Peacat search (http://microbiology.se/sw/petkit) against all assembled contigs (Bengtsson-Palme et al. 2014). The resulting consensus sequences were tested for circularity using Pemap (http://microbiology.se/software/petkit/), but no evidence of circularity was found.

We mapped the annotated mitochondrial genomes of *C. hyperboreus* and *C. sinicus* to the longest contig obtained from the assembly and were able to identify all expected mitochondrial genes.

The mitochondrial sequence of *C. glacialis* is 20,674 bp long and contains 13 protein-coding genes (total of 3458 amino acids), 2 rRNA genes, 22 tRNA genes and 1 putative control region. The sequence is composed of 31.7% base A, 28.8% base T, 19.6% base C and 19.9% base G. Ribosomal 12S and 16S RNA are 656 bp and 1138 bp long, respectively. The sequence has been deposited in GenBank under the accession number MF422146.

A phylogenetic analysis was performed using all coding genes for 18 species of Crustaceans (including 10 copepods species) and two hexapods as out-groups (*Japyx solifugus* and *Campodea fragilis)*. The phylogenetic tree was reconstructed with a maximum likelihood method using PHYML (Guindon and Gascuel 2003) (GRT + I + G model, 1000 bootstraps) (Figure 1). All copepods formed a monophyletic group and *C. glacialis* clustered with the other two *Calanus* mitogenome (100% support). Given the ecological importance of *C. glacialis* within the Arctic ecosystem, the newly determined mitogenome will be useful for investigating the history of *C. glacialis* populations and their spatiotemporal variability.

CONTACT Marvin Choquet marvin.choquet@nord.no Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

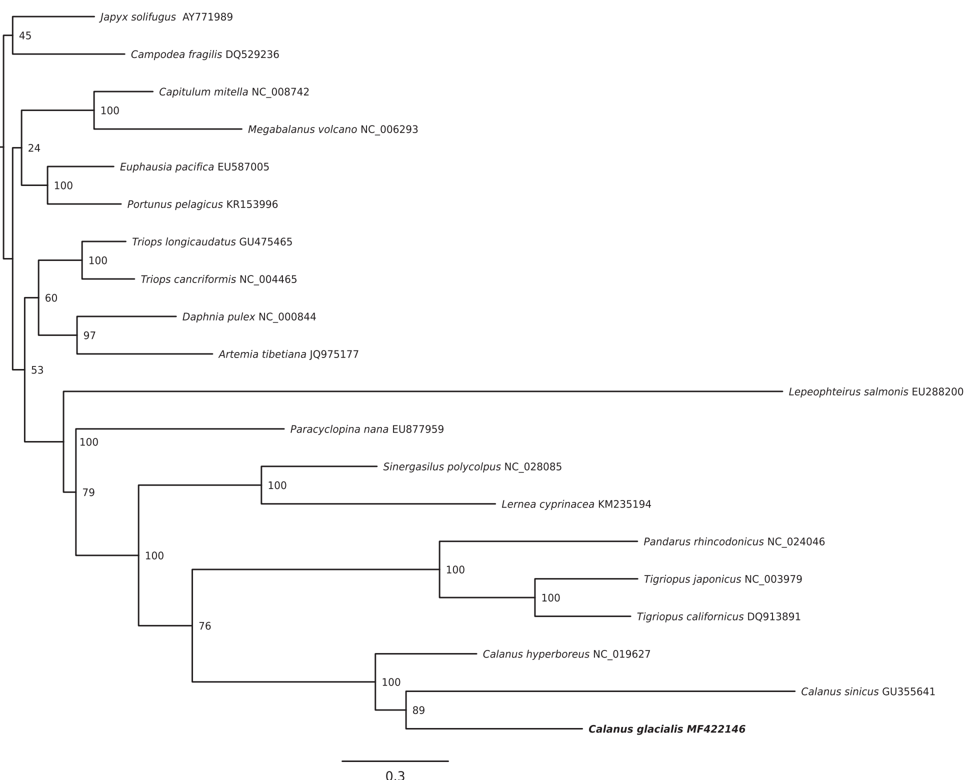


Figure 1. Phylogenetic tree of *C. glacialis* and 19 other species, with *Japyx solifugus* and *Campodea fragilis* as out-groups. ML bootstrap values (1000 replications) are indicated in front of each node.

Downloaded by [81.191.247.64] at 16:33 07 August 2017

Acknowledgements

We would like to acknowledge Morten Krogstad for sampling, Anusha Dhanasiri and Martina Kopp for technical support in the lab.

Disclosure statement

The authors have no conflict of interest to declare.

Funding

This project was supported by the Norwegian Research Council (216578) and Nord University.

ORCID

Marvin Choquet http://orcid.org/0000-0001-6719-2332

References

Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman D. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Bengtsson-Palme J, Boulund F, Fick J, Kristiansson E, Larsson DJ. 2014. Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India. Front Microbiol. 5:648.

Boisvert S, Laviolette F, Corbeil J. 2010. Ray: simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J Comput Biol. 17:1519–1533.

Falk-Petersen S, Hopkins C, Sargent J. 1990. Trophic relationships in the pelagic, Arctic food web. Trophic relationships in the marine environment. Aberdeen: Aberdeen University Press; p. 315–333.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 52:696–704.

Kim S, Lim BJ, Min GS, Choi HG. 2013. The complete mitochondrial genome of Arctic *Calanus hyperboreus* (Copepoda, Calanoida) reveals characteristic patterns in calanoid mitochondrial genome. Gene. 520:64–72.

Minxiao W, Song S, Chaolun L, Xin S. 2011. Distinctive mitochondrial genome of Calanoid copepod *Calanus sinicus* with multiple large non-coding regions and reshuffled gene order: useful molecular markers for phylogenetic and population studies. BMC Genomics. 12:73.

Smolina I, Kollias S, Poortvliet M, Nielsen TG, Lindeque P, Castellani C, Moller EF, Blanco-Bercial L, Hoarau G. 2014. Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. Mol Ecol Resour. 14:1072–1079.

WoRMS Editorial Board. 2017. World Register of Marine Species. [accessed 2017 Jul 31]. Available from http://www.marinespecies.org at VLIZ.

Paper V

Population genomics of zooplankton: how to deal with the challenges of large genomes in non-model species

Marvin Choquet[1*] - Irina Smolina[1] (shared first authorship), Anusha K. S. Dhanasiri[1], Martina Kopp[1], Alexander Jueterbock[1], Arvind Y. M. Sundaram[2], Galice Hoarau[1]

[1] *Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway*
[2] *Norwegian Sequencing Centre, Department of Medical Genetics, Oslo University Hospital, Oslo, Norway*

*Corresponding author: Marvin Choquet
marvin.choquet@nord.no

Abstract

Zooplankton organisms are key ecological links in marine food webs and are often used as climate change indicators. Nonetheless, genomics resources are still very scarce for most groups. This lack of genomic resources together with the complexity and large sizes of genomes associated have hampered the development of zooplankton population genomics. We used the species *Calanus finmarchicus*, ecologically important in the North Atlantic, as a model, we experimented with two methods of genome reduction, ddRAD and capture enrichment, in order to develop an effective method to characterized a relatively high number of SNPs. *C. finmarchicus* has a particularly large and complex genome. ddRAD sequencing was the least promising, limited by the genome size of *C. finmarchicus*. We obtained 1,871 SNPs, but only 343 SNPs were in common among the locations tested. Contrastingly, the capture enrichment method, based on a set of 2,656 unique genes showed very promising results, with about 140k SNPs in total, of which 55k were in common among the three populations investigated. In addition, we tested the capture probes on the congeneric species *C. glacialis*, and obtained about 115k SNPs, with 77k in common between the 3 populations tested. We further recommend a simplified workflow based on our experience in capture methodology to obtain tens of thousands of SNPs when working with non-model species with large and complex genomes.

Introduction

Assessment of population genetic metrics for non-model species and in particular marine zooplankton has been usually limited to a small number of loci, most frequently mitochondrial genes (Kelly and Palumbi, 2010, Peijnenburg and Goetze, 2013), that do not reflect genome-wide diversity and differentiation (Morin et al., 2004). Recent technological advances in Next Generation Sequencing (NGS) have dramatically increased sequencing throughput, reduced costs, and – together with the development of bioinformatic tools – opened new possibilities for the development of novel genomic resources in any species. NGS has enabled the genotyping of thousands of loci dispersed throughout the genome, leading to fast development of population genomics (Davey et al., 2011). Nevertheless, whole-genome sequencing for many individuals of species with genomes larger than 1 Gb remains hampered by cost and bioinformatics challenges associated with the volume of data (Davey et al., 2011, Narum et al., 2013). However, as many biological questions can be answered with only a subset of the genome, genome reduction sequencing methods have become increasingly popular. Genome reduction methods include various protocols that can be broadly grouped into amplicon, transcriptome, restriction digest, and capture enrichment sequencing (Crawford and Oleksiak, 2016, McCormack et al., 2013, Schlötterer et al., 2014). These methods not only have allowed the analysis of 1000s of Single Nucleotide Polymorphism variants (SNPs) in many individuals (McCormack et al., 2013) but also resulted in higher coverage per locus and increased accuracy of polymorphism detection (Ekblom and Galindo, 2011).

Restriction site associated DNA sequencing (RAD-seq, Baird et al., 2008; ddRAD-seq, Peterson et al., 2012) has been most commonly used for non-model organisms, due to several advantages: no need for existing genomic resources, no allele-specific expression bias, no species-specific reagents, low costs, and genotyping of SNPs from anonymous loci throughout the genome (Davey et al., 2011, McCormack et al., 2013, Schlötterer et al., 2014). RAD-seq involves an enzymatic digestion of DNA and selected sequencing of fragments flanked by recognition sites of restriction enzymes (Baird et al., 2008). The double digest RAD-seq, a modification of RAD-seq more suitable for species with large genomes, utilizes a double enzymatic digestion of DNA and allows to adjust the number of fragments to be sequenced via the choice of restriction enzymes and the size selection of digested fragments (Peterson et al., 2012).

Although costs of DNA sequencing and library preparation continue to decrease, genotyping many samples for large-scale population genetic studies is still expensive. To overcome this problem, sequencing of the pooled DNA of individuals belonging to same species or population (Pool-seq) has been proposed as an alternative to

sequencing of individually barcoded individuals (Futschik and Schlötterer, 2010). Pool-seq is cost- and time-effective, and reliable inferences of population parameters and allele frequencies can be obtained by following recommendations for statistical analysis of pooled samples (Futschik and Schlötterer, 2010, Gautier et al., 2013, Schlötterer et al., 2014). Pool-ddRAD-seq can be especially beneficial for small organisms, such as copepods, for which the DNA yield from a single individual is limited.

Another genome reduction method becoming more and more popular is sequence capture enrichment. Different strategies of capture exist and have been reviewed by Mamanova et al. (2010). The method consists in capturing specific fragments of the genome by hybridization with probes that contain complementary sequences of the targets (Gnirke et al., 2009, Jones and Good, 2016). However, the initial requirement of sequence capture is to have prior knowledge of the sequences targeted in order to design a capture probe set (Elshire et al., 2011). As this can represent a real challenge in the case of non-model species, strategies have been developed such as using an assembled transcriptome as reference instead of a genome (Bi et al., 2012). A transcriptome is much easier to obtain than a genome, especially for species with a large genome. Since capture enrichment method requires genomic resources for capture probe design, it has been less popular in non-model species. Nevertheless, the method offers valuable advantages such as the possibility to use the capture probe set on closely related species (Hancock-Hanser et al., 2013, Hedtke et al., 2013, Lemmon et al., 2012, Vallender, 2011) and on historical and degraded DNA (Carpenter et al., 2013, Enk et al., 2014, Mason et al., 2011). In addition, compared to other methods of genome reduction, sequence capture enrichment requires very little DNA for library preparation and it usually produces better data quality, more consistent loci coverage, and subsequently, more accurate SNP calling (Gnirke et al., 2009, Harvey et al., 2013, Ku et al., 2012, Tewhey et al., 2009).

In the present study, we experimented with these two methods of genome reduction to investigate the population genomics of the ecologically important planktonic marine copepod *Calanus finmarchicus*. This copepod is known to have a surprisingly large genome (6.34 Gbp; McLaren et al., 1988), and due to this it remains a non-model organism for genetics despite its paramount ecological importance. Indeed, *Calanus finmarchicus* plays an important role in linking lower and higher trophic levels (reviewed by Falk-Petersen et al., 2009) and dominates the mesozooplankton assemblage of the North Atlantic in terms of biomass (Head et al., 2003). Population genetic structure and connectivity of this key species have been long-standing subjects of research, reflecting the history of genetic marker development from allozymes

(Sywula et al., 1993) and mitochondrial genes (Bucklin et al., 1996, Bucklin and Kocher, 1996) to microsatellites (Provan et al., 2009) and few nuclear SNPs (Unal and Bucklin, 2010). All studies have suggested high levels of polymorphism and gene flow, however conclusions have ranged from lack of population genetic structure based on six microsatellites (Provan et al., 2009) to a large-scale structure based on 24 SNPs in three nuclear genes (Unal and Bucklin, 2010). The question of *C. finmarchicus* population differentiation within populations in the North Atlantic remains open and requires a NGS approach that can yield thousands of genetic markers throughout genome.

Our primary goal was to find the most effective genome reduction method that would allow us to get a high number of independent SNPs to conduct a proper population genomics study. The large and complex genome of *C. finmarchicus* and the lack of existing genome resources for this non-model species were the main challenges we had to face. We started with a ddRAD-seq approach on pooled individual, which resulted in an insufficient number of SNPs. We then switched to a capture-enrichment approach, and used an available transcriptome reference to build genomic reference for development of the final set of baits. The present paper is not aimed to compare the two approaches, but rather to describe the challenges of each method related to non-model species with large genome, and what are the options to obtain a sufficient number of SNPs. We also tested a potential of cross-species capture hybridisation on the closely related *C. glacialis*. Out of this experience, we further propose a simplified method to reach similar results when working on non-model species with large genome.

Materials and methods

1- Samples and DNA extraction

Zooplankton was sampled from 9 locations that span the distributional range of *Calanus finmarchicus* (Table 1). Samples were collected by vertical tows between either 0-100 m or 0-200 m depth using WP2 or similar nets with mesh size of 200 µm and preserved in 95 % undenatured ethanol, with subsequent change of ethanol after 24 h. Total genomic DNA was extracted individually using the E.Z.N.A. Insect DNA Kit (Omega Bio-Tek) or E.Z.N.A. DNA/RNA Isolation Kit (Omega Bio-Tek) according to manufacturer's instruction. Species identification was confirmed using a set of six nuclear insertion-deletion markers (InDels) (Smolina *et al.* 2014).

2- Pool ddRAD-seq

2.1 Library preparation and sequencing

Sixteen individuals of *C. finmarchicus* per location were used for further analyses (Table 1) since genetic differentiation can be accurately estimated from few individuals (n ≤ 6) if number of markers is large (> 100) (Willing et al., 2012). Pool-ddRAD-seq libraries were prepared according to Peterson *et al.* (2012), with minor modifications. Given the level of multiplexing and sequencing effort for this study, we aimed for approximately 8,000 fragments per pool, with coverage 200 X. Two restriction enzymes, one with rare and another with common recognition sites, were selected based on *in silico* digestion of a previously-sequenced portion of the *C. finmarchicus* genome (≈ 0.5%) (Smolina et al., 2014) using the R package SimRAD (Lepais and Weir, 2014). Among six commonly used restriction enzymes (*Msp*I, *Bam*HI, *Eco*RI, *Sbf*I, *Sph*I, *Nla*III, *Mlu*CI), the pair *Sbf*I-*Eco*RI resulted in a suitable number of fragments in the 400-500 bp size range. DNA from the 16 individuals from each sampling location was pooled in equal amounts. The six pools of approximately 100 ng were individually digested overnight at 37°C with 20,000 units of *Eco*RI-HF (New England BioLabs) and *Sbf*I-HF (New England BioLabs) enzymes in CutSmart buffer (New England BioLabs) with total volume of 50 µl. Reactions were cleaned with the Agencourt AMPure XP system (Beckman Coulter) using 1.5x volume of the AMPure reagent. Digested DNA fragments were ligated with adapters for 10 min at room temperature using the Quick Ligation Kit (New England BioLabs). The ligation products were cleaned with Agencourt AMPure XP system (Beckman Coulter), as described above. Ligated and cleaned fragments in the range 480–580 bp were selected separately for each library using 2% agarose gel E-Gel SizeSelect (Invitrogen). All obtained DNA was PCR amplified for 15 cycles (annealing temperature of 62°C) in total volume of 50 µl using the Phusion High-Fidelity PCR kit (New England BioLabs) and according to the manufacturer´s instructions. Reactions were cleaned with the Agencourt AMPure XP system (Beckman Coulter) using 0.8x volume of the reagent. Products were size selected using 2% agarose gel E-Gel SizeSelect (Invitrogen) and ran on the Agilent 2200 TapeStation System (Agilent Technologies) for quantification. Libraries were pooled in equal amounts and prepared for sequencing with MiSeq Reagent Kits v2 on a 500 cycles chip (Illumina).

2.2 SNPs genotyping

Sequenced reads were identified by six barcodes using DDemux (Rašić et al., 2014). Adapter and quality trimming was performed using Trim Galore! (Babraham Bioinformatics), with Phred quality > 20 and minimum length of 20 bp. As there is no reference genome for *Calanus*, trimmed reads were *de novo* assembled with the ddRAD assembler Rainbow v2.0.3 (Chong et al., 2012) and clustered with CD-HIT (Fu et al., 2012), as implemented in a special pipeline for ddRAD population genomics, dDocent (Puritz et al., 2014).

To take into account specificity of Pool-seq data, the data were further analysed according to recommendations of Schlötterer *et al.* (2014). Cleaned reads were mapped to the *de novo* assembly using Bowtie2 v2.2.3 (Langmead and Salzberg, 2012), with the following parameters: end-to-end -I 40 -X 850 -N 1 -L 20 -D 20 -R 3 -i S,1,0.50. Only pairs of reads that were uniquely and concordantly mapped with quality ≥ 20 were selected for further analysis using a custom-made script. To avoid false-positive SNPs, reads were realigned around InDels using the Genome Analysis Toolkit (DePristo et al., 2011), as described in Wit *et al.* (2012). SNPs were called on a merged bam file containing aligned reads of all six locations with SNVerGUI (Wang et al., 2012), which calls single nucleotide variants from pooled data and evaluates the significance of a candidate locus to be a variant. SNPs were filtered from obtained polymorphic sites with the following settings: minimum number of alleles = 2 and minimum coverage = 96 per pool (i.e., 6x per individual) using VCFtools (Danecek et al., 2011). Only high-quality SNPs were passed into further analysis. The regions with coverage > 1000x were excluded, since they may represent clusters of multi-copy genetic regions that can inflate the number of false-positive SNPs.

3- Sequence capture enrichment

3.1 Development of genomic reference for *C. finmarchicus*

As there is so far no good quality genomic reference for *C. finmarchicus*, we used the transcriptome from Lenz et al. (2014) to design a probe set to capture gene sequences and assemble them in a custom genomic reference. We selected all the sequences that were at least 749 bp long (= 29,518 sequences). To these core sequences, we added 38 unique transcript that are involved in thermal stress response of *C. finmarchicus* (Smolina et al., 2015). This resulted in a total of 29,556 sequences. We blasted (blastn in Geneious v9.1.8) each transcript against the whole transcriptome and kept only

unique sequences in order to reduce false-positive SNPs from paralogous and repeated regions. We then trimmed these sequences to the first 200 bp in order to keep mostly the 5'UTR regions since they should be particularly enriched in SNPs (Schork et al., 2013). The resulting 18,588 sequences of 200 bp lengths (representing a total size of 3,717,600 bp) were sent to a manufacturing company (Roche NimbleGen©) to produce 120-mer probes.

The sequencing library was prepared according to manufacturer's recommendations (*NimbleGen SeqCap EZ Library SR version 4.2*) using a single individual of *C. finmarchicus.* Starting DNA amount of 100 ng was fragmented to an average size of 500 bp by sonication using a Covaris® shearing instrument. Fragmented DNA was subsequently end-repaired and A-tailed using the KAPA library preparation kit (Kapa Biosystems) following the producer's recommendations. Index adapter (SeqCap Adapter Kit, Roche NimbleGen©) was ligated at 16°C overnight to allow better efficiency. Size-selection was performed using AMPure XP beads (Beckman Coulter) to keep only the fragments larger than ca. 450 bp. Further, fragments were amplified by PCR for 7 cycles using KAPA HiFi HotStart ReadyMix (Kapa Biosystems) and cleaned-up with AMPure XP beads (Beckman Coulter). Final size selection of fragments with lengths in the range of 450 – 700 bp was performed using Pippin Prep (Sage Science). Hybridisation of the sample to the probes was performed overnight, with the SeqCap Hybridization and Wash Kit (Roche NimbleGen©). Capture of the DNA by streptadivin-coated magnetic beads was done in 45 minutes in a thermocycler at 47°C. The captured sequences were cleaned up (AMPure XP beads – Beckman Coulter) and then amplified by PCR for 10 cycles. The resulting amplified sequences were cleaned up using AMPure XP beads (Beckman Coulter). Sequencing was carried out on a sequencer MiSeq (Illumina) with v.3 chemistry, in paired-end mode, using a 2 x 300 bp sequencing kit.

To check if generated reads could be used directly for SNPs discovery, we mapped raw reads to the 29,556 full-length transcriptomic contigs initially used for the capture probe design using BWA-MEM tool in default mode (Li, 2013). The mapping success was quite low (only 28% of the reads uniquely mapped with high quality score), thus strengthening the requirement for a truly genomic reference. Therefore, raw reads were filtered to remove duplicates and low complexity sequences using PRINSEQ (Schmieder and Edwards, 2011) and then assembled using the MaSuRCA assembler v.3.2.2 (Zimin et al., 2013).

3.2 Design of capture probe set

(Illustrated in Fig. 1)

From the genomic data generated by the previous sequencing, we tried to identify all the transcripts that had been successfully captured and sequenced. For that purpose, we started by downloading *C. finmarchicus* transcriptomic reads available on NCBI (https://www.ncbi.nlm.nih.gov/ - Ref: PRJNA236528) and mapped them to the 29,556 full-length transcriptomic sequences using Bowtie2 v2.2.3 (Langmead and Salzberg, 2012). Then, to identify targeted genes that were successfully captured and sequenced, the 33,294,898 reads that mapped to the selected transcripts were mapped to the MaSuRCA assembly of genomic data using TopHat RNAseq splice aware mapper v2.1.1 (Trapnell et al., 2009). This resulted in 9,225,593 reads that were mapped to 36,223 contigs. Corresponding consensus sequences were generated and the 36,223 contigs were blasted (blastn in Geneious v9.1.8) against themselves, in order to keep only single-copy genes, resulting in 3,500 contigs with only 1 hit. We performed the second blast of 3,500 contigs against the full MaSuRCA assembly, and selected the 2,223 contigs with 1 hit plus 433 other contigs blasting more than once but with 97% or more pairwise identity. We finally obtained a total of 2,656 contigs with length from 302 to 3,029 bp. The longest sequences were trimmed to a maximum length of 1,500 bp. The set of 2,656 sequences with length from 302 to 1,500 bp was used for development of a new capture probe set (in total 2,106,591 bp) by the MYcroarray® MYbaits® company with 80 mer probes and a 2x tiling density.

3.3 Library preparation and sequencing: standard protocol

A total of 48 libraries were prepared with the NEXTflex™ Rapid Pre-Capture Combo Kit (Bioo Scientific©), following the producer's protocol v15.07. We prepared 24 libraries of *C. finmarchicus* individuals and 24 libraries of *C. glacialis* individuals, including 3 populations for each species with 8 individuals per population (Table 1). Fragmentation of DNA by sonication was carried out on 150 ng of DNA per individual using Covaris® instrument (Covaris) and aiming for a fragment size of ca. 300 bp. For each of the 48 libraries, end-repair and adenylation reactions were performed according to the manufacturer protocol. Single adapter indexing was performed for each library overnight at 16°C. Ligated fragments were cleaned up with AMPure XP beads (Beckman Coulter). Fragments were amplified by PCR for 7 cycles. All libraries were then multiplexed by pooling 20 ng from each. The pool was size selected using Pippin Prep technology (Sage Science) with a range size of 400-550 bp. A total amount of 120 ng of genomic DNA was hybridised to the probes at 65°C during 3 nights. Then,

the capture reaction was conducted using DYNAbeads® MyOne™ Streptadivin C1 beads to bind the hybridised targets during 30 min at 65°C. The captured DNA was amplified by PCR for 8 cycles using KAPA® HiFi HotStart ReadyMix (Kapa Biosystems), but as no output was visualized on an Agilent 2200 TapeStation system (Agilent Technologies), we ran 10 supplementary PCR cycles. A final cleaning with AMPure XP beads was performed after PCR. Sequencing was run on the NextSeq sequencing platform (Illumina), using 4.5% Phix control and 2x150 bp mid-output kit.

To quickly evaluate if the method has resulted in sufficient outcome we trimmed demultiplexed reads using TrimGalore! (Babraham Bioinformatics) with settings minimum length > 50 bp and Phred score > 20. Then, trimmed reads were mapped to the targeted 2,656 contigs from the MaSuRCA assembly using Bowtie2 v2.2.3 (Langmead and Salzberg, 2012) with parameters –N 1 –L 20 –D 20 –R 3 -i S,1,0.50. Only uniquely and properly mapped reads without duplicates were used to assess number of reads on target (within targeted 2,656 contigs) following technical note from Roche NimbleGen (http://sequencing.roche.com/content/dam/rochesequence/worldwide/resources/NG_SeqCap_TchNote_EvalEpiData.pdf). The evaluation showed very low high quality mapping success (6.9 % on average), high level of duplication (7 fold), and low per cent of reads on target (1 %) with average coverage of 5.6 X. Therefore, we tried to optimise the protocol of library preparation and capture.

3.4 Library preparation and sequencing: optimised protocol

A total of 36 libraries were prepared, including 24 libraries for *C. finmarchicus*, with 8 individuals per populations for 3 populations; and 12 libraries for *C. glacialis* including 6 individuals from one population, and 3 individuals per populations for two other populations. The number of individuals per pool for subsequent capture reaction was voluntarily different between the 2 species. *C. finmarchicus* has an estimated genome size of 6.5 Gb (haploid), and *C. glacialis* has an estimated genome size of 12 Gb (McLaren et al., 1988), therefore by reducing number of individuals of *C. glacialis* in the final pool, we ensured that the same DNA amount will represent sufficient number of full genome copies. As previously, each DNA libraries were fragmented by sonication (starting DNA amount = 60 ng for each *C. finmarchicus* libraries and 130 ng for each *C. glacialis* libraries), and prepared using the NEXTflex™ Rapid Pre-Capture Combo Kit (Bioo Scientific©) as described previously until the hybridisation step. Another modification of the protocol included two rounds of hybridisation with lower temperature. Libraries were pooled per species and hybridisation reactions were

performed separately for the two species, using only 4 µL of beads (instead of 5.5 µL) for each reaction. To allow more specificity, the temperature for hybridization was set to 60°C. After the first round of capture, libraries were amplified by PCR for 8 cycles. The second round of hybridisation was conducted with the remaining 1.5 µL of beads, for each of the 2 pools (*C. finmarchicus* and *C. glacialis*), and was followed by a second capture, and 6 more cycles of post-capture PCR. These two rounds of hybridisation and capture are supposed to increase the efficiency of the capture, and to help to capture more sequences on-target. Finally, the 2 pools were mixed together in equal proportions, with 4.5% Phix control and sequenced on the NextSeq platform (Illumina), with a 2x150 bp mid-output kit.

The resulting raw sequences were demultiplexed and mapped directly to the full MaSuRCA assembly using BWA-MEM (Li, 2013). Generated metrics showed a global satisfying percentage of high quality reads mapping back to the reference (with 36% of *C. finmarchicus* reads in average mapping properly, uniquely, and without duplicates, while 22 % of *C. glacialis* reads in average mapped properly, uniquely and without duplicates). Hence, we decided to continue the analyses with SNP calling.

3.5 SNPs genotyping

After mapping with BWA-MEM, we only kept the reads that mapped to the reference uniquely, concordantly and in proper pairs. We removed the duplicates using Picard tools (http://broadinstitute.github.io/picard), and realigned alignments around InDels using the suite of tools GATK v3.8-0 (DePristo et al., 2011). Then we called variants using the walker HaplotypeCaller in GATK v3.8-0 (Van der Auwera et al., 2013). Obtained vcf file was filtered to remove InDels and keep only bi-allelic SNPs using VCFtools (Danecek et al., 2011). Then, the vcf file was subset into two vcf files containing SNPs of each species separately. For *C. finmarchicus*, SNPs with coverage below 5 X or above 132 X were filtered out. For *C. glacialis*, SNPs with coverage below 5 X or above 212 X were filtered out. Choice of appropriate coverage per site was based on the average SNP coverage and standard deviation across all SNPs in each species: higher threshold = [average+2*standard deviation], and the lower threshold was defined as 5X for both species. Numbers of SNPs in common between species and between populations were calculated using BCFtools v1.6.

Results

1- Pool ddRAD-seq

On average, 99.85 % of the reads passed quality filtering. *De novo* assembly resulted in 41,500 contigs covering 17,886,794 bp and with a mean GC content of 38.9 %. Mean length of contigs was 430.9 bp, with a range of 259 – 758 bp. However, when the forward and reverse reads did not overlap, they were connected with up to 10 "N" bases. In total, 2,836 contigs were annotated (6.83 %). Most (87.23 %) of the reads mapped in proper pairs and with a mapping quality > 20. However, on average, only 27.56 % of the reads mapped uniquely (Fig. 2), resulting in the mean coverage per location library of 16 X (Fig. 3a).

Overall, in all six libraries, 24,701 single nucleotide variants were detected with SNVerGUI software. Among these variants, 15,285 were high-quality SNPs, but only 1,871 SNPs had minimum coverage of 96 X per library (Fig. 3b). With maximum and minimum coverage thresholds, the average number of SNPs per location was 510, with a minimum of 211 in the Bay St-Lawrence (due to fewer sequenced reads) and a maximum of 625 in the Norwegian Sea. SNPs were distributed on contigs with up to 14 SNPs on one same contig (Fig. 4). The SNPs were distributed over 99 contigs on average (Table 2) per library. A total of 343 SNPs, located over 32 contigs, were found to be common among all six populations.

2- Sequence capture enrichment

The capture probe set based on genomic reference generated from a first transcriptome-based capture experiment was applied to 48 individuals (24 *C. finmarchicus* and 24 from 3 populations each – see Table 1). Following the standard manufacturer's protocol we obtained ca. 1.5 million reads per individuals, for which only 7% mapped properly to the reference. After removal of duplicates, only 1% of the reads were left, and considered as on target. From ca. 1.5 million reads per individuals, we ended up with ca. 30 000 reads per individuals, which did not allow for enough coverage for SNP calling.

The optimised version of the protocol used on 36 individuals (24 *C. finmarchicus* and 12 *C. glacialis* from 3 populations each – see Table 1) yielded much better results. We obtained in average 4.5 million reads per individuals for *C. finmarchicus*, and 17.5 million reads for *C. glacialis* in average per individuals (Table 3). Mapping was more

successful than previously with an average of 1.6 million reads properly and uniquely mapping to the reference (after duplicates removal) for *C. finmarchicus*. This represented in average 36% of the initial number of reads sequenced per individual, with a range from 31.4% to 41.6% depending on individuals (Table 3). For *C. glacialis*, 3.8 million reads mapped properly and uniquely to the reference in average per individual (after duplicates removal). This represents 22% of the initial number of reads sequenced in average per individual, with a range from 20.1% to 24.3% (Table 3).

From these remaining reads, variant calling resulted in 140,859 high quality and sufficiently covered SNPs in total for *C. finmarchicus*, ranging from 85,882 to 96,102 SNPs per population (Table 4). These SNPs were distributed across 4,603 contigs (Fig. 5a). A total of 54,848 SNPs were in common among all 3 populations (Table 4). For *C. glacialis*, 115,928 high quality and sufficiently covered SNPs were called, ranging from 87,829 to 104,623 SNPs per population (Table 4). These SNPs were distributed across 5,363 contigs (Fig. 5b). A total of 77,402 SNPs were in common among all 3 populations (Table 4). Interestingly, 31,283 SNPs were found in common between *C. finmarchicus* and *C. glacialis* (Table 4).

Discussion

The goal of this study was to experiment with genome reduction methods in order to obtain a sufficient amount of SNPs to conduct a robust population structure study on *Calanus finmarchicus*, zooplankton non-model species with large and complex genome.

Our analytical approach using Pool-ddRAD-seq data and strict quality filtering resulted in an average of 510 SNPs per location, with 343 SNPs in common among all locations. The number of SNPs detected in this study is much lower than for most RAD-seq or Pool-seq, studies, for which numbers of SNPs are typically thousands (e.g. Campana et al., 2015, Hohenlohe et al., 2011, Pukk et al., 2015, Reitzel et al., 2013). Low SNP numbers are likely mainly due to the low coverage of ddRAD contigs (mean coverage was 16 X per pool), which does not allow calling SNPs with high confidence. The low contig coverage is likely due to the *in silico* under-estimation of ddRAD fragments and low success (28 %) of unique and high-quality mapping, which are related to the large and complex *C. finmarchicus* genome. Further, mapping of reads from the pooled samples is challenging, since the population pool may contain high levels of polymorphism; however too-liberal mapping parameters increase the chances

of incorrect mapping of reads (Kofler et al., 2011). Therefore, we followed recommendations for the pooled data (Kofler et al., 2011, Schlötterer et al., 2014), resulting in fewer SNPs, but higher confidence levels.

Most commonly, species that are analysed using RAD- or ddRAD-seq have genome sizes < 5 Gb, resulting in high numbers of usable fragments and SNPs (e.g. DaCosta and Sorenson, 2014, Davey et al., 2011, Pukk et al., 2015). While few studies have investigated species with large genomes for applications other than marker development, these serve to highlight the difficulties (e.g. Deagle et al., 2015). Subsequently, ddRAD-seq was designed to allow more flexible control over the number of obtained contigs and can result in several orders of magnitude variation in the number of fragments by using restriction enzymes and selecting fragments of specific sizes (Baird et al., 2008, Peterson et al., 2012), which is beneficial for analysis of species with large genomes. However, selection of the enzyme pair can be difficult in uncharacterized genomes, and may differ among species with similar genome sizes and GC composition, due to different frequencies of the restriction sites (Davey et al., 2011). Furthermore, if only a small fraction of a species' genome is known, it may not be fully representative of the entire genome. In the present study, the selected enzyme pair resulted in 5 times more ddRAD contigs than expected, and consequently reduced the average coverage per contig and the number of contigs suitable for further analysis. Thus, transfer of the enzyme pair suitable for *C. finmarchicus* to other *Calanus* species may be problematical without pilot studies, particularly due to genome size variation among the species, with *C. finmarchicus* having the smallest genome (McLaren et al., 1988). Overall, this study and recent similar ones (e.g. Deagle et al., 2015, Pukk et al., 2015) suggest that significant challenges remain for use of both, Pool-seq and (dd)RAD-seq, for species with relatively large genomes (> 5 Gb).

For the targeted *C. finmarchicus*, more than 140 k SNPs were found, with about one third of them in common between the populations and thus usable for population structure analysis. Even for the congeneric species *C. glacialis*, on which we also experimented the capture, based on *C. finmarchicus* design, we still obtained over 115k SNPs with more than 77k in common among populations. It should be noted that these SNPs were located in ca. 5000 contigs and thus only represent ca. 5000 independent loci. Nonetheless, this is a dramatic change of scale compared to what was obtained before from the RAD-seq approach. Furthermore the physical proximity of many of the SNPs detected opens up the possibility to infer the precise sequence (phase) of alleles on each homologous copy of a chromosome (Delaneau et al., 2013, Snyder et al., 2015). Such phased haplotype can then be used to infer ancestry and demographic history (Song et al., 2016) or detect selection. The generated set of SNPs

represents an invaluable tool to assess genetic variability among *C. finmarchicus* and *C. glacialis* populations.

As zooplankton organisms represent a key link in marine food webs and have a crucial role in marine ecosystems, the understanding of the functioning of their populations is critical especially given they are often used as beacons of climate changes (Hays et al., 2005, Richardson, 2008). So far, challenges linked to the incredibly common large genomes size of these organisms have hampered gene flow studies (reviewed by Bucklin et al., unpubl.), and therefore the knowledge on this essential aspect of zooplankton is very limited. The methods we tested on *C. finmarchicus*, typical non-model species with large genome, showed that RAD-seq is probably not the most adequate for such organisms. In contrast, sequence capture enrichment represents a very promising approach, and seems particularly suitable for zooplankton organisms given that the starting amount of DNA can be even lower than 10 ng (Chung et al., 2016). As the sequence capture enrichment technic usually requires a reference, which is rarely available for zooplankton organisms, the approach we followed offers a reasonable compromise (See Fig. 1). However, with the constant reduction in cost of sequencing, such approach could be further simplified by generating genomic data directly by shotgun sequencing and aligning together genomic and transcriptomic sequences in order to target mainly genic or anonymous intergenic regions, depending on the purpose of the study (Fig. 1).

Acknowledgements

We would like to warmly acknowledge Ann Bucklin for the samples she provided, and for her comments on the manuscript. We thank Maja Hatlebakk and Morten Krogstad for collecting samples used in the capture-enrichment part of this study. We thank Torkel Gissel Nielsen (Danish Technical University) for collection of samples around Greenland, and Ebru Unal (University of Connecticut and Mystic Aquarium) for sorting of samples from the Gulf of St. Lawrence. Special thanks to the captains, crews and scientists who participated in the research cruises providing samples used for this study: RV Johan Hjort cruise 201208 and RV GO Sars cruise 201305. This study was funded by the Norwegian Research Council (HAVKYST 216578) and EU-FP7 Eurobasin (Grant agreement 264933).

References

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. & Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one,* 3**:** e3376.

Bi, K., Vanderpool, D., Singhal, S., Linderoth, T., Moritz, C. & Good, J.M. (2012). Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC genomics,* 13**:** 403.

Bucklin, A. & Kocher, T.D. (1996). Source regions for recruitment of *Calanus finmarchicus* to Georges Bank: evidence from molecular population genetic analysis of mtDNA. *Deep Sea Research Part II: Topical Studies in Oceanography,* 43**:** 1665-1681.

Bucklin, A., Sundt, R.C. & Dahle, G. (1996). The population genetics of *Calanus finmarchicus* in the North Atlantic. *Ophelia,* 44**:** 29-45.

Campana, M.G., Robles García, N.M. & Tuross, N. (2015). America's red gold: multiple lineages of cultivated cochineal in Mexico. *Ecology and evolution,* 5**:** 607-617.

Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillén, S. & Nekhrizov, G. (2013). Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *The American Journal of Human Genetics,* 93**:** 852-864.

Chong, Z., Ruan, J. & Wu, C.-I. (2012). Rainbow: an integrated tool for efficient clustering and assembling RAD-seq reads. *Bioinformatics,* 28**:** 2732-2737.

Chung, J., Son, D.-S., Jeon, H.-J., Kim, K.-M., Park, G., Ryu, G.H., Park, W.-Y. & Park, D. (2016). The minimal amount of starting DNA for Agilent's hybrid capture-based targeted massively parallel sequencing. *Scientific reports,* 6**:** 26732.

Crawford, D.L. & Oleksiak, M.F. (2016). Ecological population genomics in the marine environment. *Briefings in functional genomics,* 15**:** 342-351.

Dacosta, J.M. & Sorenson, M.D. (2014). Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS One,* 9**:** e106713. DOI 10.1371/journal.pone.0106713

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

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. & Blaxter, M.L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature reviews Genetics,* 12**:** 499.

Deagle, B.E., Faux, C., Kawaguchi, S., Meyer, B. & Jarman, S.N. (2015). Antarctic krill population genomics: apparent panmixia, but genome complexity and large population size muddy the water. *Molecular ecology,* 24**:** 4943-4959.

Delaneau, O., Howie, B., Cox, A.J., Zagury, J.-F. & Marchini, J. (2013). Haplotype estimation using sequencing reads. *The American Journal of Human Genetics,* 93**:** 687-696.

Depristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., Del Angel, G., Rivas, M.A. & Hanna, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics,* 43**:** 491-498.

Ekblom, R. & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity,* 107**:** 1.

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

Enk, J.M., Devault, A.M., Kuch, M., Murgha, Y.E., Rouillard, J.-M. & Poinar, H.N. (2014). Ancient whole genome enrichment using baits built from modern DNA. *Molecular Biology and Evolution,* 31**:** 1292-1294.

Falk-Petersen, S., Mayzaud, P., Kattner, G. & Sargent, J.R. (2009). Lipids and life strategy of Arctic *Calanus*. *Marine Biology Research,* 5**:** 18-39.

Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics,* 28**:** 3150-3152.

Futschik, A. & Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics,* 186**:** 207-218.

Gautier, M., Foucaud, J., Gharbi, K., Cézard, T., Galan, M., Loiseau, A., Thomson, M., Pudlo, P., Kerdelhué, C. & Estoup, A. (2013). Estimation of population allele frequencies from next‐generation sequencing data: pool‐versus individual‐based genotyping. *Molecular Ecology,* 22**:** 3766-3779.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., Leproust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S. & Russ, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology,* 27**:** 182-189.

Hancock‐Hanser, B.L., Frey, A., Leslie, M.S., Dutton, P.H., Archer, F.I. & Morin, P.A. (2013). Targeted multiplex next‐generation sequencing: advances in techniques of mitochondrial and nuclear DNA sequencing for population genomics. *Molecular ecology resources,* 13**:** 254-268.

Harvey, M., Smith, B., Glenn, T., Faircloth, B. & Brumfield, R. (2013). Sequence capture versus restriction site associated DNA sequencing for phylogeography. arXiv preprint. *arXiv preprint arXiv:13126439*.

Hays, G.C., Richardson, A.J. & Robinson, C. (2005). Climate change and marine plankton. *Trends in ecology & evolution,* 20**:** 337-344.

Head, E., Harris, L. & Yashayaev, I. (2003). Distributions of *Calanus* spp. and other mesozooplankton in the Labrador Sea in relation to hydrography in spring and summer (1995–2000). *Progress in Oceanography,* 59**:** 1-30.

Hedtke, S.M., Morgan, M.J., Cannatella, D.C. & Hillis, D.M. (2013). Targeted enrichment: maximizing orthologous gene comparisons across deep evolutionary time. *PLoS One,* 8**:** e67908.

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

Jones, M.R. & Good, J.M. (2016). Targeted capture in evolutionary and ecological genomics. *Molecular Ecology,* 25**:** 185-202.

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

Kofler, R., Orozco-Terwengel, P., De Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C. & Schlötterer, C. (2011). PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. *PloS one,* 6**:** e15925.

Ku, C.-S., Wu, M., Cooper, D.N., Naidoo, N., Pawitan, Y., Pang, B., Iacopetta, B. & Soong, R. (2012). Exome versus transcriptome sequencing in identifying coding region variants. *Expert review of molecular diagnostics,* 12**:** 241-251.

Langmead, B. & Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods,* 9**:** 357-359.

Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic biology,* 61**:** 727-744.

Lenz, P.H., Roncalli, V., Hassett, R.P., Wu, L.S., Cieslak, M.C., Hartline, D.K. & Christie, A.E. (2014). De novo assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda)--the dominant zooplankter of the North Atlantic Ocean. *PLoS One,* 9**:** e88589. DOI 10.1371/journal.pone.0088589.

Lepais, O. & Weir, J.T. (2014). SimRAD: an R package for simulation ‐ based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. *Molecular ecology resources,* 14**:** 1314-1321.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:13033997*.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J. & Turner, D.J. (2010). Target-enrichment strategies for next-generation sequencing. *Nature methods,* 7**:** 111-118.

Mason, V.C., Li, G., Helgen, K.M. & Murphy, W.J. (2011). Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Research,* 21**:** 1695-1704.

Mccormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution,* 66**:** 526-538.

Mclaren, I., Sevigny, J.-M. & Corkett, C. (1988). Body sizes, development rates, and genome sizes among *Calanus* species. *Biology of Copepods.* Springer.

Morin, P.A., Luikart, G. & Wayne, R.K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution,* 19**:** 208-216.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R. & Hohenlohe, P.A. (2013). Genotyping ‐ by ‐ sequencing in ecological and conservation genomics. *Molecular ecology,* 22**:** 2841-2847.

Peijnenburg, K.T. & Goetze, E. (2013). High evolutionary potential of marine zooplankton. *Ecology and Evolution,* 3**:** 2765-2781.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One,* 7**:** e37135. DOI 10.1371/journal.pone.0037135.

Provan, J., Beatty, G.E., Keating, S.L., Maggs, C.A. & Savidge, G. (2009). High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus* finmarchicus. *Proceedings Biological sciences / The Royal Society,* 276**:** 301-7. DOI 10.1098/rspb.2008.1062.

Pukk, L., Ahmad, F., Hasan, S., Kisand, V., Gross, R. & Vasemägi, A. (2015). Less is more: extreme genome complexity reduction with ddRAD using Ion Torrent semiconductor technology. *Molecular ecology resources,* 15**:** 1145-1152.

Puritz, J.B., Hollenbeck, C.M. & Gold, J.R. (2014). dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ,* 2**:** e431.

Rašić, G., Filipović, I., Weeks, A.R. & Hoffmann, A.A. (2014). Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, Aedes aegypti. *BMC genomics,* 15**:** 275.

Reitzel, A., Herrera, S., Layden, M., Martindale, M. & Shank, T. (2013). Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular ecology,* 22**:** 2953-2970.

Richardson, A.J. (2008). In hot water: zooplankton and climate change. *Ices Journal of Marine Science,* 65**:** 279-295. DOI 10.1093/icesjms/fsn028.

Schlötterer, C., Tobler, R., Kofler, R. & Nolte, V. (2014). Sequencing pools of individuals--mining genome-wide polymorphism data without big funding. *Nature reviews Genetics,* 15**:** 749.

Schmieder, R. & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics,* 27**:** 863-864.

Schork, A.J., Thompson, W.K., Pham, P., Torkamani, A., Roddey, J.C., Sullivan, P.F., Kelsoe, J.R., O'donovan, M.C., Furberg, H. & Schork, N.J. (2013). All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. *PLoS genetics,* 9**:** e1003449.

Smolina, I., Kollias, S., Møller, E.F., Lindeque, P., Sundaram, A.Y., Fernandes, J.M. & Hoarau, G. (2015). Contrasting transcriptome response to thermal stress in two key zooplankton species, *Calanus finmarchicus* and *C. glacialis*. *Marine Ecology Progress Series,* 534**:** 79-93.

Smolina, I., Kollias, S., Poortvliet, M., Nielsen, T.G., Lindeque, P., Castellani, C., Moller, E.F., Blanco-Bercial, L. & Hoarau, G. (2014). Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular ecology resources,* 14**:** 1072-9. DOI 10.1111/1755-0998.12241.

Snyder, M.W., Adey, A., Kitzman, J.O. & Shendure, J. (2015). Haplotype-resolved genome sequencing: experimental methods and applications. *Nature Reviews Genetics,* 16**:** 344-358.

Song, S., Sliwerska, E., Emery, S. & Kidd, J.M. (2016). Modeling human population separation history using physically phased genomes. *Genetics***:** genetics. 116.192963.

Sywula, T., Glazewska, I., Koszteyn, J., Kwasniewski, S. & Sell, J. (1993). An analysis of the population structure of *Calanus* cf. *finmarchicus* (Copepoda) from the Hornsund fjord region, Spitsbergen. *Var Evol 2,* 3**:** 113-119.

Tewhey, R., Nakano, M., Wang, X., Pabón-Peña, C., Novak, B., Giuffre, A., Lin, E., Happe, S., Roberts, D.N. & Leproust, E.M. (2009). Enrichment of sequencing targets from the human genome by solution hybridization. *Genome biology,* 10**:** R116.

Trapnell, C., Pachter, L. & Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics,* 25**:** 1105-1111.

Unal, E. & Bucklin, A. (2010). Basin-scale population genetic structure of the planktonic copepod *Calanus finmarchicus* in the North Atlantic Ocean. *Progress in Oceanography,* 87**:** 175-185. DOI 10.1016/j.pocean.2010.09.017.

Vallender, E.J. (2011). Expanding whole exome resequencing into non-human primates. *Genome biology,* 12**:** R87.

Van Der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy‐Moonshine, A., Jordan, T., Shakir, K., Roazen, D. & Thibault, J. (2013). From FastQ data to high‐confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics***:** 11.10. 1-11.10. 33.

Wang, W., Hu, W., Hou, F., Hu, P. & Wei, Z. (2012). SNVerGUI: a desktop tool for variant analysis of next-generation sequencing data. *Journal of medical genetics***:** jmedgenet-2012-101001.

Willing, E.-M., Dreyer, C. & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS One,* 7**:** e42649.

Wit, P., Pespeni, M.H., Ladner, J.T., Barshis, D.J., Seneca, F., Jaris, H., Therkildsen, N.O., Morikawa, M. & Palumbi, S.R. (2012). The simple fool's guide to population genomics via RNA‐Seq: an introduction to high‐throughput sequencing data analysis. *Molecular ecology resources,* 12**:** 1058-1067.

Zimin, A.V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S.L. & Yorke, J.A. (2013). The MaSuRCA genome assembler. *Bioinformatics,* 29**:** 2669-2677.

-

Tables

Table 1. *Calanus finmarchicus* and *C. glacialis* sample information.

Location	Study	Species	n	Collection date	Lat.	Long.
West Greenland (WG)	ddRAD-seq	*C. finmarchicus*	16	25.05.2012	69.23	-53.38
East Greenland (EG)	ddRAD-seq	*C. finmarchicus*	16	10.08.2012	74.31	-20.25
Barents Sea (BAR)	ddRAD-seq	*C. finmarchicus*	16	06.08.2012	70.50	19.99
Norwegian Sea (N12)	ddRAD-seq	*C. finmarchicus*	16	03.08.2012	64.67	0.00
Norwegian Sea (N13)	ddRAD-seq	*C. finmarchicus*	16	05.05.2013	65.05	-0.86
Bay St. Lawrence (STL)	ddRAD-seq	*C. finmarchicus*	16	17.04.2013	47.27	-59.80
Barents Sea	Transcriptomic Capture	*C. finmarchicus*	1	06.08.2012	70.50	19.99
Isfjord	Genomic Capture	*C. finmarchicus*	8	05.06.2016	78.32	15.15
		C. glacialis	3			
Skjerstadfjord	Genomic Capture	*C. finmarchicus*	8	26.02.2016	60.72	5.10
		C. glacialis	6			
Lurefjord	Genomic Capture	*C. finmarchicus*	8	22.06.2016	67.18	15.43
		C. glacialis	3			

Table 2. Summary of discovered SNPs using ddRAD at six locations for *Calanus finmarchicus*.

Location	Total no. SNPs	Number of unique SNPs	Number of contigs with SNPs
WG	570	147	96
EG	559	148	116
BAR	565	162	109
N12	625	193	124
N13	535	165	106
STL	211	34	43

Table 3. Capture enrichment method efficiency for *Calanus finmarchicus* and *C. glacialis*. From initial total number of reads obtained from sequencing, to number of reads that properly mapped, uniquely and with high quality score to the reference, after duplicates removal.

	Total number of reads	Number of reads properly mapped	% reads properly mapped
Calanus finmarchicus			
CF_Is_1	4346219	1633598	37,59
CF_Is_2	4377473	1710440	39,07
CF_Is_3	3795041	1413854	37,26
CF_Is_4	5300843	2069892	39,05
CF_Is_5	6079302	2384972	39,23
CF_Is_6	5369728	2088164	38,89
CF_Is_7	4847206	2013872	41,55
CF_Is_8	2796841	1107884	39,61
CF_Lure_17	2171800	799004	36,79
CF_Lure_18	1384749	478180	34,53
CF_Lure_19	3710734	1305160	35,17
CF_Lure_20	2485355	801722	32,26
CF_Lure_21	3155247	990962	31,41
CF_Lure_22	2918617	942954	32,31
CF_Lure_23	1313948	482804	36,74
CF_Lure_24	3667871	1417456	38,65
CF_Skj_33	3910873	1358892	34,75
CF_Skj_34	3595077	1352758	37,63
CF_Skj_35	3168606	1080116	34,09
CF_Skj_36	9422745	3209456	34,06
CF_Skj_37	8644477	2838838	32,84
CF_Skj_38	7123122	2310916	32,44
CF_Skj_39	6569504	2240762	34,11
CF_Skj_40	7329317	2752352	37,55
Average			36%
Calanus glacialis			
CG_Is_10	14397845	2894612	20,1
CG_Is_11	8080749	1961926	24,28
CG_Is_16	5460660	1265816	23,18
CG_Lure_28	5344982	1284384	24,03
CG_Lure_29	28994596	6637518	22,89
CG_Lure_32	21520844	4645090	21,58
CG_Skj_43	19272833	3881326	20,14
CG_Skj_44	21781295	4749712	21,81
CG_Skj_45	21305311	4727152	22,19
CG_Skj_46	19778639	4165368	21,06
CG_Skj_47	20067631	4216582	21,01
CG_Skj_48	24591690	5396430	21,94
Average			22%

Table 4. Summary of discovered SNPs using capture enrichment

Location	Species			
	C. finmarchicus		***C. glacialis***	
	n indiv.	Total nb SNPs	n indiv.	Total nb SNPs
Isfjord	8	93 834	3	87 829
Skjerstadfjord	8	96 102	6	104 623
Lurefjord	8	85 882	3	92 837
SNPs per species	24	140 859	12	115 928
SNPs in common among 3 populations	54 848		77 402	
SNPs in common between species	31 283			

Figures

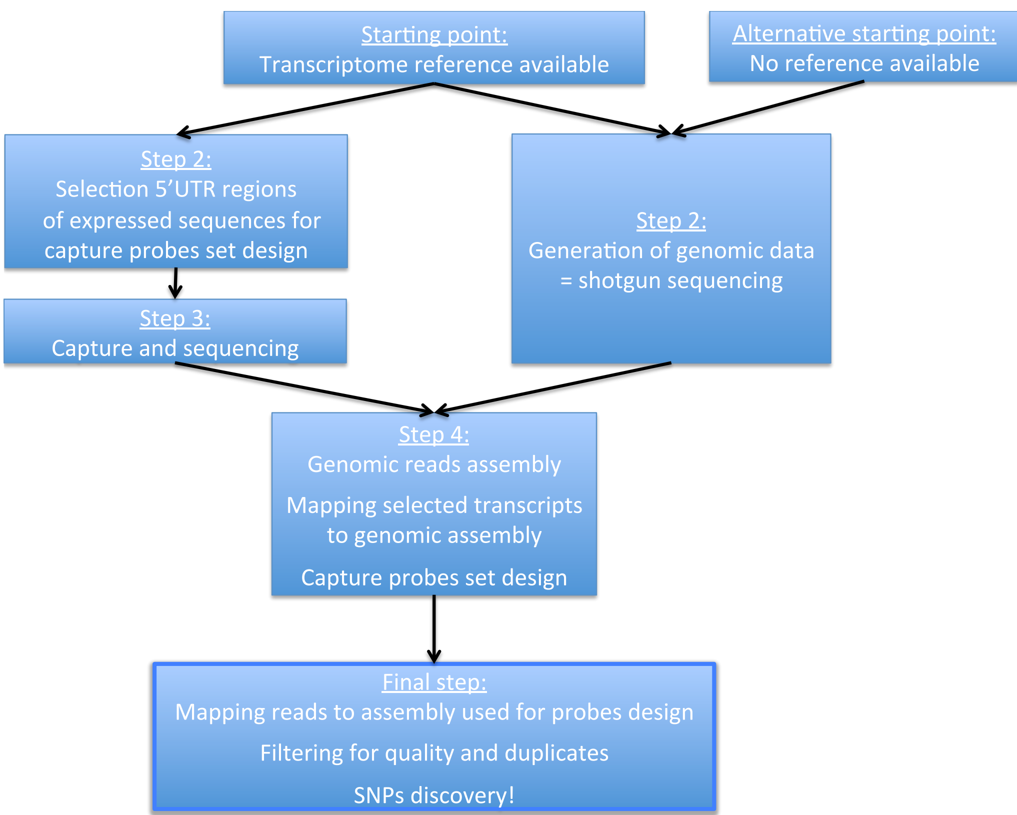


Fig. 1. Suggested workflow for SNP discovery in non-model species. Step 2: on the left – protocol described here, on the right: recommended simplified version.

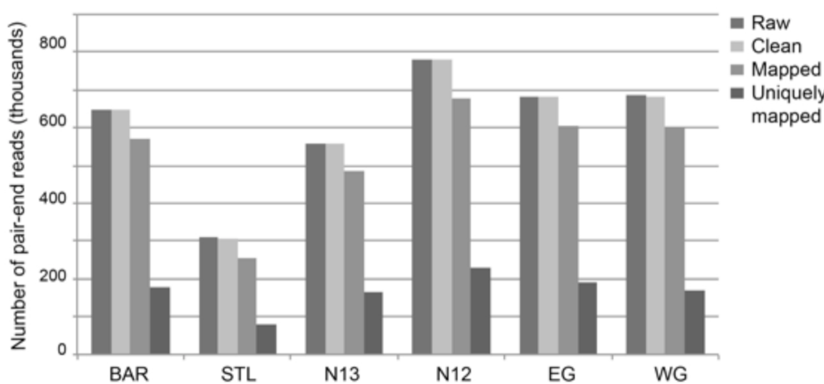


Fig. 2. Number of Illumina paired-end reads for each library throughout bioinformatics analysis. Raw, total sequenced reads; Clean, reads without adapters, a Phred quality > 20, longer than 20 bp and trimmed adapters; Mapped, reads mapped to the assembly with Phred quality > 20 and in proper pairs; Uniquely mapped, reads that mapped to unique site.

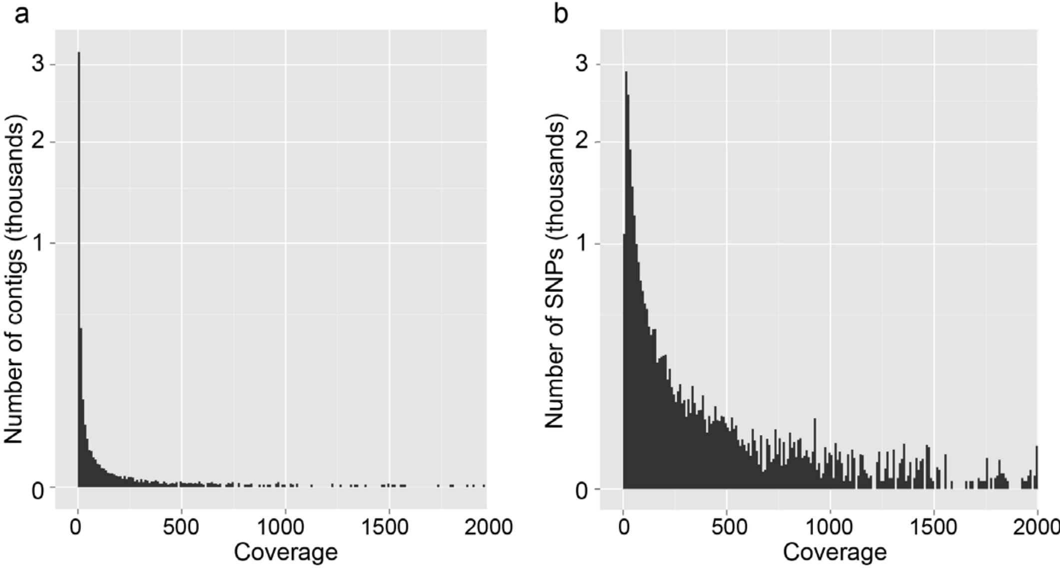


Fig. 3. Number of contigs (a) and SNPs (b) in relation to depth of coverage for ddRAD experiment.

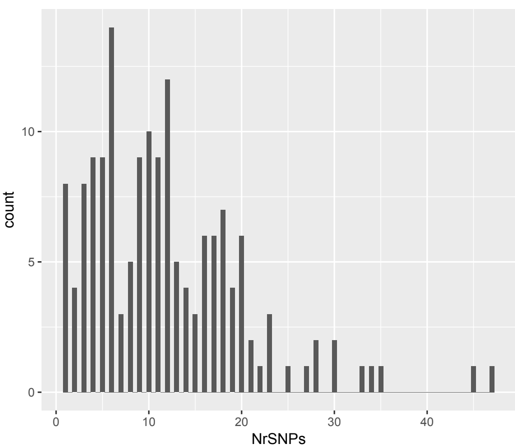


Fig. 4. Frequency of contigs according to number of SNPs per contig obtained from Pool-ddRAD-seq in *Calanus finmarchicus*.

a)

b)

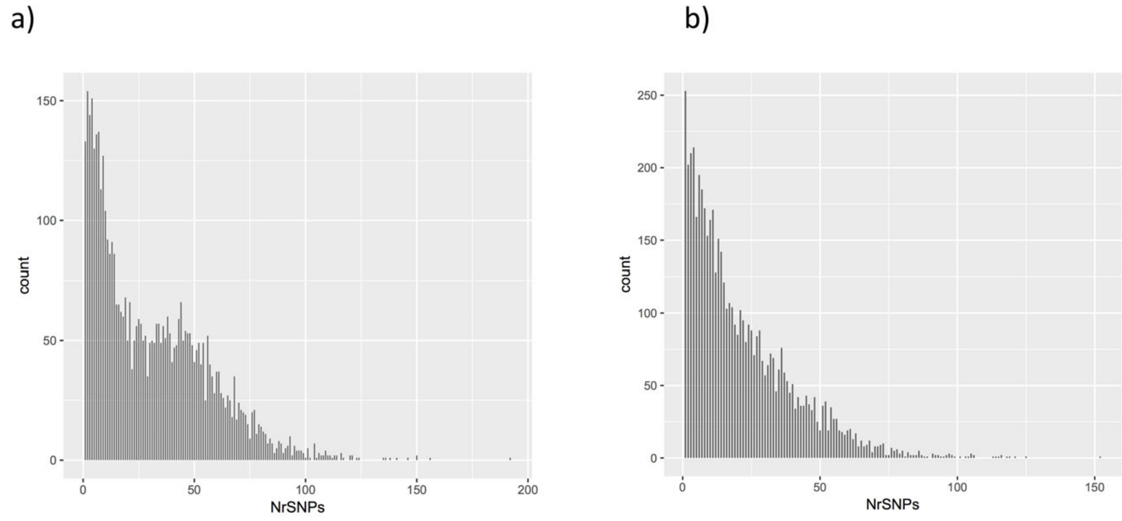


Fig. 5. Frequency of contigs according to number of SNPs per contig obtained from targeted capture enrichment in a) *Calanus finmarchicus* and b) *Calanus glacialis*.

Annex 1

Population Genomics of Marine Zooplankton

Chapter 19 in: *Population Genomics: Marine Organisms*

Editors: Om P. Rajora and Marjorie Oleksiak

Ann Bucklin (Corresponding author)
Department of Marine Sciences, University of Connecticut
1080 Shennecossett Road, Groton, CT 06340 USA
Tel. 860-405-9260; Fax 860-405-9153
Email: ann.bucklin@uconn.edu

Kate DiVito
Dept. of Molecular and Cell Biology, University of Connecticut
354 Mansfield Rd (U-1131), Storrs, CT 06269 USA
Email: kate.divito@uconn.edu

Irina Smolina
Marine Ecology Research Group
Faculty of Biosciences and Aquaculture, Nord University
Universitetsalléen 11, 8026 Bodø, Norway
Email: irina.smolina@nord.no

Marvin Choquet
Marine Ecology Research Group
Faculty of Biosciences and Aquaculture, Nord University
Universitetsalléen 11, 8026 Bodø, Norway
Tel: +47 75 51 77 56; Fax: +47 94 89 08 53
Email: marvin.choquet@nord.no

Jennifer M. Questel
Department of Marine Sciences, University of Connecticut
1080 Shennecossett Road, Groton, CT 06340 USA
Tel. 860-405-9255; Fax 860-405-9153
Email: jennifer.questel@uconn.edu

Galice Hoarau
Marine Ecology Research Group
Faculty of Biosciences and Aquaculture, Nord University
Universitetsalléen 11, 8026 Bodø, Norway
Tel: +47 75 51 74 97
Email: galice.g.hoarau@nord.no

Rachel J. O'Neill
Center for Genome Innovation and Institute for Systems Genomics
Dept. of Molecular and Cell Biology, University of Connecticut
354 Mansfield Rd (U-1131)
Storrs, CT 06269 USA
Tel: 860-486-6031; Fax: 860-486-1936
Email: rachel.oneill@uconn.edu

Table of Contents

19.1. Abstract

19.2. Introduction
19.2.1. Introduction to population genomics
19.2.2. Introduction to marine zooplankton
    19.2.2.1. Biodiversity
    19.2.2.2. Biogeography
    19.2.2.3. Life history
19.2.3. Genomic resources for marine zooplankton
    19.2.3.1. Published genomic resources
    19.2.3.2. Genome sizes of zooplankton
    19.2.3.3. Mitogenomes
    19.2.3.4. Transcriptomic resources

19.3. Applications of population genomics for marine zooplankton
19.3.1. Population genetic diversity and structure
19.3.2. From population genetics to population genomics
19.3.3. Genomic basis of adaptation
19.3.4. Metagenetics and metabarcoding

19.4. Case studies of marine zooplankton
19.4.1. *Calanus finmarchicus* (Copepoda)
19.4.2. *Centropages typicus* (Copepoda)
19.4.3. *Tigriopus californicus* (Copepoda)
19.4.4. *Acartia tonsa* (Copepoda)
19.4.5. *Euphausia superba* (Euphausiacea)
19.4.6. *Meganyctiphanes norvegica* (Euphausiacea)
19.4.7. *Pleurobrachia bachei* (Ctenophora)
19.4.8. *Spadella cephaloptera* (Chaetognatha)
19.4.9. *Salpa thompsoni* (Tunicata)

19.5. Present-day challenges and future opportunities
19.5.1. Additional genomic resources for marine zooplankton species
19.5.2. Sampling zooplankton in the global ocean
19.5.3. Species identification
19.5.4. Sequencing small-size organisms
19.5.5. Genomic basis of adaptation

19.6. Acknowledgements

19.7. Table Legends

19.8. Figure Legends

19.9. Literature Cited

Population Genomics of Marine Zooplankton

Ann Bucklin, Kate DiVito, Irina Smolina, Marvin Choquet, Jennifer M. Questel, Galice Hoarau, and Rachel J. O'Neill

19.1. Abstract

The exceptionally large population size and cosmopolitan biogeographic distribution that distinguish many – but not all – marine zooplankton species generate similarly exceptional patterns of population genetic and genomic diversity and structure. The phylogenetic diversity of zooplankton has slowed the application of population genomic approaches, due to lack of genomic resources for closely-related species and diversity of genomic architecture, including highly-replicated genomes of many crustaceans. Use of numerous genomic markers, especially single nucleotide polymorphisms (SNPs), is transforming our ability to analyze population genetics and connectivity of marine zooplankton, and providing new understanding and different answers than earlier analyses, which typically used mitochondrial DNA and microsatellite markers. Population genomic approaches have confirmed that, despite high dispersal potential, many zooplankton species exhibit genetic structuring among geographic populations, especially at large ocean-basin scales, and have revealed patterns and pathways of population connectivity that do not always track ocean circulation. Genomic and transcriptomic resources are critically needed to allow further examination of micro-evolution and local adaptation, including identification of genes that show evidence of selection. These new tools will also enable further examination of the significance of small-scale genetic heterogeneity of marine zooplankton, to discriminate genetic "noise" in large and patchy populations from local adaptation to environmental conditions and change.

Keywords: Zooplankton, Population genomics, Transcriptomics, Evolution, Population genetics

19.2. Introduction

II.A. Introduction to population genomics

Population genomic approaches entail simultaneous sampling of numerous variable loci within a genome and allow inference of locus-specific effects (Baird et al. 2008). These powerful new techniques are transforming our understanding of the population genetics, connectivity, demographic history, and local adaptation of marine organisms (Crawford and Oleksiak 2016; Pogson 2016). Genotyping hundreds to thousands of genetic markers for multiple individuals across populations or species has enabled identification of selectively-neutral markers that can be used for a wide variety of analyses (Luikart et al. 2003; Baird et al. 2008). Discrimination of statistical 'outlier' loci allows examination of the impacts of selection and evidence of local adaptation (Stapley et al. 2010). Whole-genome analysis of non-model organisms has enabled new insights into underlying evolutionary forces. However, significant challenges remain for whole-genome analysis of non-model organisms, thus necessitating and encouraging broad use of approaches that require little or no prior genomic data. These include reduced-representation genomic DNA libraries (Reitzel et al. 2013), genotyping-by-sequencing (Elshire et al. 2011), and exon-capture (Hodges et al. 2007; De Wit et al. 2015; Jones and Good 2016), although the latter requires prior knowledge of gene architecture. In broad view, population genomic approaches have enormous potential to yield significant new understanding of the ecological and evolutionary dynamics of zooplankton and other marine organisms.

19.2.1. Introduction to marine zooplankton

<u>19.2.2.1. Biodiversity</u>: The marine zooplankton assemblage includes ~6,000 described species of holoplanktonic metazoan organisms that complete their entire cycle in the water column (Wiebe et al. 2010). The phylogenetic diversity of this assemblage is impressive, with 11 phyla and 27 orders represented (Bucklin et al. 2010b). However, these numbers most likely markedly underestimate the actual biodiversity – perhaps by several orders of magnitude – due to the presence of cryptic variation within geographically widespread species or sibling species swarms, as well as undiscovered species in under-sampled or explored habitats (Bucklin et al. 2010a; Beaugrand 2017). Molecular approaches, including DNA barcoding and metabarcoding, are providing important new insights into this 'hidden diversity' of marine zooplankton (Lindeque et al. 2013; Bucklin et al. 2016).

<u>19.2.2.2. Biogeography</u>: Global patterns of zooplankton biogeographic distributions have been well-characterized for the epipelagic (0 – 200 m) zone (Longhurst 2007). The many classical studies form a basis for ongoing examination of climate-driven range changes and regime shifts (deYoung et al. 2008). In contrast, the deep ocean, including the mesopelagic (200 – 1,000 m) and bathypelagic (1,000 – 4,000 m), remains under-sampled and poorly-known (but see Wiebe et al. 2010; Laakmann et al. 2012). Many species exhibit cosmopolitan distributions, with ranges spanning multiple ocean basins and broad latitudinal ranges (Peijnenburg and Goetze 2013). However, there are many exceptions to this oversimplified description, likely resulting from specific habitat requirements, restricted gene flow, or relict

populations (Chust et al. 2016). Further complicating analysis of species distributional patterns are rather characteristic high ratios of local-to-global species diversity; a net sample from oceanic waters may contain hundreds of species of copepods or ~10% of the global total (Kuriyama and Nishida 2006).

19.2.2.3. Life history: Many zooplankton species have life histories entailing multiple stages with different micro-habitat preferences and requirements. Some exhibit alternation of sexual and asexual generations. Most are relatively short-lived organisms, with generation spans from several months to a couple of years. As a group, marine zooplankton are useful indicators of impacts of environmental variability or climate change, since they are rapid-responders in terms of species distribution and abundance. The exceptional diversity of marine zooplankton – in terms of phylogenetic biodiversity, pelagic biogeography, and life history variation – provided a unique opportunity to examine ecological and evolutionary genomic responses. This review will summarize new knowledge resulting from population genomic examination of the genetic diversity and structure, phylogeography and connectivity, demographic history, and local adaptation of marine metazoan holozooplankton.

19.2.3. Genomic resources for marine zooplankton

19.2.3.1. Published genomic resources: It can be argued that there are no universally-accepted model species among the marine zooplankton; in many cases, there are no closely-related model organisms to which extrapolations or comparisons can be made (Ellegren 2014). However, the number of marine zooplankton species targeted for genome-scale studies is growing, including species ranging phylogenetically from the Cnidaria to the Urochordata and including ecologically-important or keystone species for some pelagic ecosystems, such as the Southern Ocean salp, *Salpa thompsoni* (Jue et al. 2016) (Table 1).

For the most part, marine zooplankton species targeted for reference sequencing and assembly have been identified by their impact to ongoing comparative genomic studies or as part of larger genome consortia. An example of this latter group is the genome sequence for the copepod *Eurytemora affinis*, a species targeted for sequencing as part of the i5K Pilot Project aimed at sequencing 28 arthropod genomes (i5K Consortium 2013; Eyun et al. 2017). Currently, assembled genomes are available for species representing only a snapshot of some of the major lineages of eukaryotes and a small sampling of the species diversity of the pelagic realm (Table 1). A significant factor in the identification of a target species for a genome assembly effort is the estimated genome size. Notably, all the reference genomes available are from organisms whose genome size estimates are significantly smaller than 1GB, presumably since the depth of coverage required is low enough to represent a feasible investment of resources in terms of fiscal and computational effort. While reference quality assemblies are ideal (e.g., *Oikopleura dioica*, Denoeud et al. 2010), lower coverage assemblies can still provide a high enough N50-value (i.e., the weighted median statistic such that 50% of the entire assembly is contained in contigs or scaffolds equal to or larger than this value) to afford extensive gene predictions (e.g., Jue et al. 2016).

Recently, mining genome databases such as NCBI and the SRA (Short Read Archive) for partial genome sequences has afforded broader comparisons among species lacking a fully

assembled genome. For example, a newly derived reference for the common estuarine copepod *E. affinis* was compared to short read genomic sequence data from two other copepods, the freshwater cyclopoid copepod, *Mesocyclops edax* (SRX246444 and SRX246445; Sun et al. 2014) and the North Atlantic copepod, *Calanus finmarchicus*) (SRX456026; Smolina et al. 2014), revealing species-specific adaptations of the chemosensory related gene families to environments (Eyun et al. 2017).

19.2.3.2. Genome size in the zooplankton: The average estimated genome sizes (haploid nuclear DNA contents) of holoplankton species are in general far above 1 GB (Fig. 1) and varies more than 900-fold, from 0.07 GB in *Oikopleura dioica* (Appendicularia) to 63.2 GB in *Ampelisca macrocephala* (Amphipoda). Variation of genome sizes in marine zooplankton is especially large within the Copepoda with > 370-fold variation among species (Leinaas et al. 2016; Madoui et al. 2017) followed by Ostracoda and Malacostraca with around 80-fold and 70-fold variation of genome size among species, respectively (Gregory 2017; Jeffery et al. 2017). To date, genome size has been investigated for 115 species of zooplankton, with poor representation of important phyla, including Chaetognatha, Cnidaria, Ctenophora, Mollusca and Chordata.

Several trends or patterns are emerging from genomic analyses of crustaceans, although only a few species have been studied to date. First, a positive relationship between genome size (C-value) and body size has been observed in copepods (McLaren et al. 1988; Wyngaard and Rasch 2000), amphipods (Hessen and Persson 2009), and ostracods (Jeffery et al. 2017). However, there is considerable variability in genome size both among species of similar body size (Gregory et al. 2000; Leinaas et al. 2016) and within species due to environmental conditions (McLaren et al. 1988; Escribano et al. 1992; Leinaas et al. 2016). Second, genome size has been associated with specific habitats and environmental conditions. Marine crustaceans are likely to have larger genomes than freshwater and terrestrial ones (Jeffery 2015; Alfsnes et al. 2017); within the marine realm, polar species tend to have larger genomes compared to temperate species (Hessen and Persson 2009; Jeffery 2015; Leinaas et al. 2016). Jeffery (Jeffery 2015) hypothesizes that such large genomes may result from the expansion of transposable elements and other repetitive elements, due to relaxed selection for rapid development or reduced constraints on body size in predictable and stable marine polar environments, compared to more fluctuating environments.

Causes and mechanisms of genome size variability and particularly expansion of genome sizes are still not known. Among eukaryotes, genome size is positively correlated with gene number, average intron size, and number of introns per genome (Elliott and Gregory 2015). The main drivers of genome size expansion are suggested to be whole-genome duplication (polyploidization) or partial duplication events and proliferation of noncoding elements (Dufresne and Jeffery 2011).

Information on genome size, genome sequence, and karyotype is sparse in marine zooplankton, limiting our understanding of genome evolution. Nevertheless, evidence from insects and crustaceans suggest that accumulation of transposable and repetitive elements may be the primary contributor to their large genome sizes (Alfsnes et al. 2017), while polyploidization is probably not the most common driver of genome evolution in zooplankton

(Gregory and Hebert 1999). For example, species of the copepod genera *Calanus* and *Pseudocalanus* exhibit quantum shifts in genome size (C-values) within each genus, but share similar chromosome complements (McLaren et al. 1989).

Partial duplication or amplification of genomic regions may be common in large genomes of zooplankton, particularly for ribosomal rDNA and protein-coding genes. Among eukaryotes, rDNA copy number correlates positively with genome size (Prokopowich et al. 2003). For species of *Calanus*, 18S rDNA gene copy number has been found to approximately double between *C. finmarchicus* (15,300 copies; 2C = 12.95 pg) and *C. glacialis* (33,500 copies; 2C = 24.20 pg; Wyngaard et al. 1995). Transcriptomic analysis has indicated the presence of multi-copy gene families originating from multiple duplications of an ancestral gene in copepods (Lenz et al. 2014; Yang et al. 2014), euphausiids (Toullec et al. 2013; Sales et al. 2017), and pteropods (Maas et al. 2015; Thabet et al. 2017).

19.2.3.3. Mitochondrial genomes: Fragments of mitochondrial DNA were among the first molecular tools used to tackle questions related to zooplankton species identification, phylogenetics, phylogeography, and population genetics. For example the cytochrome oxidase sub-unit I is preferentially used as a barcode for metazoan (Schindel and Miller 2005), and has been used frequently for marine zooplankton (Bucklin et al. 2007, 2010a, 2011; Blanco-Bercial et al. 2014).

Recent technological advances are allowing routine sequencing of whole mitochondrial genomes (mitogenomes), with marked increase in the power of phylogenetic and phylogeographic analyses compared to use of short mtDNA sequences. Applications such as shotgun sequencing on genomic DNA using high throughput sequencing technologies afford opportunities to capture other genomes that may be resident within a sample, such as mitochondrial DNA. Given the smaller target genome size (12-20KB), mitogenomes are easier to subsample from larger datasets or to assemble using a PCR-build approach (Maricic et al. 2010; Hahn et al. 2013; Kollias et al. 2015).

Mitogenomics is a promising field of research that will contribute new insights into the phylogenetic history and evolution of planktonic species. For example, sequencing the mitogenome of the chaetognath, *Spadella cephaloptera*, allowed resolution of the phylogenetic position of the chaetognaths within Protostome lineages (Papillon et al. 2004). Only a few mitogenomes have been published thus far – especially when the species diversity of zooplankton is considered – and within those, unexpected features appear to be more common than previously thought. Mitogenomes are publicly available for a number of ecologically-important species representing diverse phylogenetic lineages of marine zooplankton (Table 2), and additional complete mitochondrial assemblies may be found within incompletely-explored genomic data. Nonetheless, the sequencing and assembly of complete mitogenomes of marine zooplankton species has progressed at a much slower pace than other for vertebrate groups (Genome 10K Community of Scientists 2009; GIGA Community of Scientists 2014).

In animals, the mitogenome is relatively well conserved, with 36 or 37 genes, including two for rRNAs, 22 for tRNAs and 12 or 13 for protein-coding genes. The mitogenomes available for marine zooplankton species indicate a general trend of high intra- and interspecific variability.

Rearrangement of gene order is exceptionally common and has been documented in amphipods (Ki et al. 2010) and ctenophores (Kohn et al. 2012), with some of the genes relocated to the nuclear genome (Pett et al. 2011). Copepods also show marked variability among congeneric species and among genera (Fig. 2; Jung et al. 2006; Minxiao et al. 2011). The most exceptional cases of mitochondrial variability documented to date are in the chaetognaths, *Spadella cephaloptera* and *Sagitta elegans*, for which natural populations exhibit unprecedented levels of intra-specific divergence (Marlétaz et al. 2017).

The variability observed in the mitogenomes of different species/lineages is also apparent in the genes content and size of these mitogenome (Table 2). The smallest mitogenome reported is the ctenophore, *Mnemiopsis leidyi*, with only 10 kb, which is missing 25 genes (Pett et al. 2011). Within the chaetognaths, mitogenomes are also very reduced compared to other metazoans, missing several common genes (Helfenbein et al. 2004; Papillon et al. 2004). On the other hand, the longest mitogenomes documented belong to the Copepods, up to 20 kb (Minxiao et al. 2011). Several mitogenomes were found to contain multiple copies of some sequences (Ogoh and Ohmiya 2004; Burton et al. 2007), or short tandem repeats, similar to microsatellites (Shen et al. 2011).

19.2.3.4. Transcriptomic resources: For some species, especially those with large, duplicated and/or evolutionarily-divergent genomes, analysis of transcriptomes has proven more feasible, accurate and cost-effective (De Wit et al. 2016). Transcriptomic data have the further advantage of allowing identification and annotation of target genes used in the examination of genomic micro-evolution and local adaptation (Havird and Santos 2016). Transcriptomic data, including partial reference transcriptomes are available for a number of marine zooplankton species (Table 3).

19.3. Applications of population genomics for marine zooplankton

19.3.1. Population genetic diversity and structure

Although many zooplankton species exhibit broad geographic distributions and appear to have high dispersal potential, both biological and physical environmental processes may limit gene flow. Previous studies have revealed significant genetic differentiation of geographic populations of marine organisms over a range of spatial scales (Hellberg 2009; Weersing and Toonen 2009). Two general principles may be gleaned from many studies of zooplankton population genetics: first, zooplankton are quite variable in many different molecular characters; second, this variability is resolved into genetically-divergent, geographically-distinct populations for only some species and at some temporal and spatial scales (Peijnenburg and Goetze 2013).

Ocean processes that are thought to be significant for population genetic structuring of zooplankton are currents, persistent eddies, ocean gyres and other physical ocean structures at the mesoscale (10s to 100s km) to large scale (100s to 1000s km). The physical structure of the ocean can alter the timing of reproduction and mortality events, providing biological barriers to gene flow. Geological features – continents, islands and other landforms,

continental shelves, seamounts, and ocean ridges – may form natural barriers to dispersal. In contrast, cosmopolitan species, which range from 40°N to 40°S and are found in every ocean basin, may have few barriers to dispersal throughout their range. These species may exhibit large-scale spatial population genetic structure due to isolation by distance (i.e., reproductive isolation resulting when the geographic range of the species far exceeds the dispersal potential of an individual).

The temporal stability of population genetic diversity and structure is an important consideration and useful metric. Since zooplankton are subject to transport in ocean currents, temporal stability of population genetic characters may indicate retention of local populations or local recruitment. An unfortunate aspect of many studies of zooplankton populations is the collection of samples from different regions during different years, thus confounding spatial and temporal variation. In relatively few studies, spatial and temporal contributions to population genetic structure have been analyzed separately using appropriately-collected samples (Goetze et al. 2015; Iacchei et al. 2017).

Patterns of genetic diversity and structure have been examined over a wide range of spatial scales for species representing many lineages of the zooplankton assemblage. Some species have been shown to be panmictic, such as *Pelagia noctiluca* (Stopar et al. 2010) and *Euphausia superba* (Deagle et al. 2015). Many species exhibit geographic variation reflecting geographic barriers and/or circulation patterns: e.g., *Tigriopus californicus* (Renaut and Dion-Côté 2016), *Eukrohnia hamata* (Kulagin et al. 2014), and *Caecosagitta macrocephala* (Miyamoto et al. 2010), to name a few. A number of species show large-scale patterns of genetic diversity associated with latitudinal gradients (e.g., Francisco et al. 2014) and among ocean basins, including *Eukrohnia hamata* (Miyamoto et al. 2012), *Pleuromama abdominalis* (Hirai and Tsuda 2015), and *Oithona similis* (Cornils et al. 2017).

The occurrence and significance of small-scale genetic patchiness in marine zooplankton populations remain a subject of study and disagreement. Such variability has been considered to reflect the genetic "noise" of large and under-sampled populations of copepods (e.g., Goetze et al. 2015). Small-scale heterogeneity was considered to reflect advective transport from diverse recruitment sources in the Antarctic krill, *Euphausia superba* (Batta-Lona et al. 2011).

Due to nearly-universal application in population genetic studies, hierarchical analysis of variance using Wright's *F*-statistics related measures (Excoffier et al. 1992) provides useful benchmarks for comparisons among species, regions, and environments. However, *F*-statistics have assumptions that are surely not met for zooplankton (Hellberg 2009), including genetic equilibrium conditions, symmetrical migration, and stable populations. The usefulness of *F*-statistics is further limited by the very large population sizes of many zooplankton, which result in relatively larger confidence intervals for very small *F* values (Waples 1998), and thus a lack of statistical significance for high gene flow species (see Waples et al. 2008). At least partly for this reason, population genetic studies of marine species have also employed various measures of oceanographic distance (Hansen and Hemmer-Hansen 2007; McGovern et al. 2010; Alberto et al. 2011; Schunter et al. 2011) and approaches such as seascape genetics (Galindo et al. 2010).

Until recently, population genetic studies have most frequently been conducted with

markers representing a very small fraction of the genome, such as individual mitochondrial or nuclear genes and microsatellites (see reviews by Avise 2009; Hellberg 2009; Peijnenburg and Goetze 2013). Rates of divergence and amounts of variation differ among these markers, but many studies have documented significant genetic differentiation of zooplankton populations at large, ocean basin scales using mitochondrial DNA (e.g., Goetze 2005; Goetze and Ohman 2010; Miyamoto et al. 2010; Blanco-Bercial et al. 2011a; Miller et al. 2012; Norton and Goetze 2013; Dawson et al. 2015) and microsatellite markers (Bolte et al. 2013; Andrews et al. 2014). A number of studies have used mitochondrial DNA markers to resolve population structure of zooplankton populations associated with physical barriers to gene flow, including ocean circulation, for copepods (Aarbakke et al. 2011; Blanco-Bercial et al. 2011b, 2014) and euphausiids (Bucklin et al. 1997; Zane et al. 1998, 2000; Zane and Patarnello 2000; Papetti et al. 2005; Patarnello et al. 2010).

Both mitochondrial and microsatellite markers continue to be widely used for population genetic analysis of zooplankton, allowing useful comparisons among diverse species and ocean environments. Studies using single markers have limitations, not least that results may differ among studies using different markers (Avise et al. 2016). In addition to their limited analytical power, studies using multiple markers can yield discordant conclusions. In particular, the haploid nature and uniparental inheritance of mitochondrial markers, and consequent smaller effective population size, may generate differences from results using nuclear markers (Toews and Brelsford 2012).

Population genomic approaches can also be used for phylogeographic analysis (i.e., the description of the geographical distributions of the genetic lineages within a population or species; Avise 2009; Avise et al. 2016). Such analysis allows for the characterization of dispersal and quantitative estimation of the rate and direction of exchange among populations. Recent reviews of larval dispersal and population connectivity (Cowen and Sponaugle 2009) and gene flow (Hellberg 2009) in the ocean have provided comprehensive assessment and analyses for marine organisms. Quantitative estimates of population persistence and directional (asymmetric) migration can also entail approaches that are less sensitive to lack of population stability and non-equilibrium conditions, typical of marine organisms (Knowles 2009). Analysis of patterns and pathways of gene flow has revealed that patterns of population connectivity of marine organisms do not always mimic major ocean currents (Kool et al. 2013; Riginos et al. 2016), even for zooplankton (Blanco-Bercial and Bucklin 2016; Questel et al. 2016).

Phylogeographic analysis can also provide a window into the evolutionary history of a population or species. Results can be interpreted to estimate and understand the age of the lineage in terms of time to coalescence (i.e., the common ancestral gene from which all current copies of the gene are descended), as well as imprints of demographic history on populations and species (Knowles 2009). Among marine zooplankton, mitochondrial markers have been used most regularly to infer demographic history (e.g., Peijnenburg et al. 2005; Aarbakke et al. 2014; Cornils et al. 2017), including marine invasions (Cristescu 2015; Lee 2016; Sherman et al. 2016), population expansions and contractions (Edmands 2001), geographic isolation giving rise to speciation events (Lee 2000; Peijnenburg et al. 2004; Miyamoto et al. 2010), and divergence of genetic lineages following major global climate events (Papadopoulos et al. 2005; Blanco-Bercial et al. 2011b; Milligan et al. 2011).

19.3.2. From population genetics to population genomics

Recent advances in High-Throughput Sequencing (HTS) have created exceptional new opportunities for analysis of population genetic diversity and structure of natural populations. Tens of thousands of genomic Single Nucleotide Polymorphisms (SNPs) can be detected and screened for use as genetic makers of population genetic diversity and structure (Helyar et al. 2011; Reitzel et al. 2013). Such population genomic approaches are being widely used among marine organisms (Bierne et al. 2016), including fishes (Hemmer-Hansen et al. 2014). In addition, HTS is yielding both deep coverage and nucleotide-level resolution in simultaneous or multiplexed analysis of numerous genes (e.g., Bybee et al. 2011). Such population genomic approaches are yielding a new view of population structure and connectivity of marine species, based on statistical discrimination of neutral, selected, and hitchhiker loci (Gagnaire and Gaggiotti 2016).

Over the last three decades, genetic research has showed continuous development and a high turnover of molecular markers, from partial DNA sequencing, restriction fragment length polymorphism (RFLP), random amplified polymorphism detection (RAPD) and amplified fragment length polymorphism (AFLP) to microsatellites, insertion-deletion polymorphism (InDel), and single nucleotide polymorphism (SNP; Schlötterer et al. 2014). Historically, development of markers was difficult and expensive for non-model organisms. However, the advent of HTS has revolutionized this by allowing the genome-wide markers in any organism and for low cost (Ekblom and Galindo 2011). Although simultaneous discovery and genotyping of genome-wide variation has become feasible for tens of individuals with small genome sizes (< 1GB), the individual sequencing of hundreds of individuals with large genomes remains prohibitively expensive (Narum et al. 2013). In addition, sequencing of the complete genome for all individuals is often unnecessary and inflates the bioinformatics demands (Narum et al. 2013). Therefore, for many studies including population genomics, it is more efficient to sequence a limited number of targeted loci, thus increasing their coverage and chance to detect true polymorphism (Ekblom and Galindo 2011).

A revolutionizing solution to address this situation was the development of genotyping-by-sequencing (GBS) approaches that allow sequencing with high throughput technology of a targeted fraction of the genome via various reduced-representation protocols (see review by Crawford and Oleksiak 2016). These approaches result in discovery and simultaneous genotyping of thousands of SNPs even in species with large genomes and little or no previous genomic information. GBS relies on various reduced-representation protocols to target a genome fraction, but four protocols are currently the most popular: RNA-seq, Ampli-seq, Cap-seq (i.e., capture enrichment), and RAD-seq (Davey and Blaxter 2010; Reitzel et al. 2013). Published reduced representation genomic resources are currently available for several species of marine zooplankton, such as the copepods, *Tigriopus californicus* (Foley et al. 2011), *Calanus finmarchicus* (Smolina 2015), and *Centropages typicus* (Blanco-Bercial and Bucklin 2016); and the euphausiid, *Euphausia superba* (Deagle et al. 2015). The number of studies using reduced representation for population genomics in marine zooplankton may be expected to expand in the near future.

The power of genomic SNPs for resolution of regional- to large-scale population structure of zooplankton has been demonstrated for several key species (see Case Studies, below). A large-

scale population genetic analysis using genomic SNPs demonstrated that RAD-seq methods performed poorly in the copepod, *Calanus finmarchicus*, which has a large and complex genome (Smolina 2015). Subsequent studies of this species using targeted resequencing (e.g., Cap-seq) showed promise for accurate SNP identification and detection of genetic structuring for this species (Choquet et al., unpubl. data). Similarly, a study of the copepod, *Centropages typicus,* by Blanco-Bercial and Bucklin (2016) using 1,000s of genomic SNPs obtained by RAD-seq revealed evidence of population structure, in contrast to an earlier study based on mitochondrial gene sequences (Castellani et al. 2012).

Genomic SNPs that show evidence of selection can provide markers of micro-evolution and local adaptation, including identification of the key genes involved in these phenomena. The use of many thousands of genomic markers will also enable further examination of the significance of small-scale genetic heterogeneity of marine zooplankton, including distinguishing genetic "noise" in large and patchy populations from local adaptation to environmental conditions. Large-scale SNP genotyping studies remain very scarce in zooplankton species, but as more studies based on these approaches are published, it will be important to resolve differing conclusions based on the various technical approaches and genetic markers employed.

19.3.3. Genomic basis of adaptation

Population genomic approaches have provided powerful new tools for detection of impacts of selection and evidence of local adaptation (Stapley et al. 2010). Patterns of variation of genomic markers can be statistically evaluated for non-neutrality and correlation with population dynamic, environmental, and evolutionary conditions and drivers (Gagnaire et al. 2015). Non-neutral markers showing evidence of selection can be used to reveal adaptation of populations to local conditions across a species range (Whitehead 2012), although other evolutionary drivers, including introgression and hitchhiking, can also cause such departures from neutrality for genomic traits (Bierne et al. 2013). Nielsen et al. (2009) concluded that few published studies have convincingly documented that non-neutral traits reflect local adaptation, citing reviews by Hedrick (2006) and Levasseur et al. (2007). Recent advances in statistical analysis of genomic markers are enabling more sensitive and accurate detection of local adaptations (Gayral et al. 2013; Savolainen et al. 2013; De Wit et al. 2015), although these are much more powerful for species with well-characterized genomes, which allows exome capture and sequencing (Jones and Good 2016).

Patterns of differential gene expression can also provide useful insights into local adaptive responses of marine organisms to environmental conditions. There are a number of such studies of marine zooplankton, including target-gene and whole-transcriptome analyses of differential gene expression patterns associated with stress responses and environmental variability (Lauritano et al. 2012; Schoville et al. 2012; De Pittà et al. 2013; Smolina et al. 2015, 2016; Roncalli et al. 2016; Batta-Lona et al. 2017). The genetic and genomic bases of such gene expression differences have received considerable attention (see review by Romero et al. 2014).

19.3.4. Metagenetics and metabarcoding

The exceptional challenge of species identification in zooplankton assemblages, resulting from both phylogenetic diversity and sibling species swarms, has encouraged the development of genetic approaches for both stand-alone and integrative use with morphological taxonomic methods (Bucklin et al. 2016). Metagenetic and metabarcoding approaches analyze DNA recovered from environmental samples and can reflect the biodiversity of entire pelagic communities (de Vargas et al. 2015), with the advantage of detecting 'hidden diversity' of marine zooplankton (Lindeque et al. 2013). These studies use 'universal' PCR primers to amplify one or more gene regions for high throughput sequencing yielding tens of millions of sequences, which are subsequently resolved into operational taxonomic units (OTUs) that can either be matched to reference databases for identification of taxa or used for various statistical measures of biodiversity (Leray and Knowlton 2016). Metabarcoding studies of marine zooplankton have ranged from analysis of the global ocean (Bik et al. 2012; de Vargas et al. 2015) to studies focused on particular habitats and ecosystems, such as estuaries (Abad et al. 2016), the Red Sea (Pearman and Irigoien 2015), among others. Challenges remain for quantitative analysis of taxa using metabarcoding, although recent studies have shown some correlation between OTU frequency and taxon biomass (Hirai et al. 2015; Sun et al. 2015).

The continuing development of sequencing technologies may soon allow a full metagenomics approach, where DNA extracted from environmental samples is sequenced and whole genomes are reconstructed from the data. These data will be invaluable resources for diverse population genomic approaches, including analysis of population genetic diversity and structure, detection of loci under selection, and genomic bases of adaptations of zooplankton species to environmental variation. Currently, both technical and bioinformatics challenges limit use of metagenomics to species with small genomes, such as the copepod, *Oithona nana* (Madoui et al. 2017).

19.4. Case studies of marine zooplankton

Population genomic approaches, entailing simultaneous sampling of numerous variable loci within a genome and the inference of locus-specific effects (Black et al. 2001; Luikart et al. 2003), are only very recently being used for analysis of marine zooplankton. Comparison between results from population genetic studies using single-markers (usually mitochondrial or microsatellite DNA) and HTS genomic markers are particularly useful to evaluate the power and precision of population genomic approaches for analysis of genetic structure, connectivity, demographic history, and local adaptation.

Several of the marine zooplankton species analyzed using population genomic approaches belong to the crustacean Subclass Copepoda, which comprises more species than any other zooplankton group, including many that are ecologically important, numerically predominant, and geographically widespread. Genomic analysis of copepods has been a focus of research, although progress has been hampered by the exceptionally large genome sizes of many species (Bron et al. 2011; Wyngaard et al. 2011; Jeffrey 2015).

19.4.1. *Calanus finmarchicus* (Copepoda): The planktonic copepod *Calanus finmarchicus* (Fig. 3) is thought to be the most abundant metazoan in the ocean; the species is ubiquitous in coastal and open ocean cold-temperate regions of the North Atlantic Ocean (Planque et al. 1997); within this area, the species may contribute >70% of total copepod biomass (Head et al. 2003) and occupies a pivotal position in ocean food webs (Falk-Petersen et al. 2007). Population genetic studies using mitochondrial DNA (e.g., Bucklin et al. 1996) and microsatellites (Provan et al. 2009) have shown high levels of gene flow and little or no significant population genetic structure at any spatial scale. Studies using SNPs in targeted gene regions suggested genetic differentiation among samples from different water masses and ocean basins (Bucklin and Kaartvedt 2000; Unal and Bucklin 2010 Fig. 4). Population genomic analyses of *C. finmarchicus* have been impeded by the large size of its genome (C-value = 6.48 pg; McLaren et al. 1988), typical of crustaceans. Smolina (2015) used a genotyping-by-sequencing approach (ddRADseq; Peterson et al. 2012) to characterize genomic SNPs in pooled samples of *C. finmarchicus* collected across the North Atlantic Ocean. Significant population differentiation was observed among locations, although the allelic nature of the SNP variants in the pooled samples could not be confirmed due to the highly-replicated genome (Smolina 2015). An ongoing study by this group is analyzing genomic SNPs in targeted gene regions to allow confirmation of allelic variation despite genome size (Choquet et al. 2017a). A partial reference transcriptome for the species (Lenz et al. 2014) is allowing evaluation of evidence of local adaptation based on transcriptomic and target gene analysis (e.g., Roncalli et al. 2016).

19.4.2. *Centropages typicus* (Copepoda): Blanco-Bercial and Bucklin (2016) used genomic SNPs detected by 2b-RADseq analysis (Wang et al. 2012) to examine population genetic structure of the copepod *Centropages typicus* (Fig. 5) in the North Atlantic Ocean. Thousands of genomic SNP markers were identified; loci showing evidence of positive selection were removed from analysis (Foll and Gaggiotti 2008). Statistical analysis of molecular variance (Excoffier and Lischer 2010) revealed significant differences between continental shelf populations of the NE and NW Atlantic populations, in contrast with an earlier study by Castellani et al. (2012), which showed no structuring using a mitochondrial COI gene region, but some differentiation of NE and NW Atlantic populations based on a nuclear rRNA internal transcribed spacer (ITS) region. Genotyping-by-sequencing (RADtag sequences) of *C. typicus* yielded 675 loci used by Blanco-Bercial and Bucklin (2016) to test hypotheses of dispersal and directional migration (Beerli 2012). Among five different gene flow models (Fig. 6), the full migration model showed the highest support. These results demonstrate the power of population genomic approaches to resolve patterns and pathways of dispersal of a high gene flow species in a dynamic and complex current system. Such analyses can also be used to examine the genomic basis of observed local adaptation of this species to environmental variability among regions or along a latitudinal gradient (Carlotti et al. 2007).

19.4.3. *Tigriopus californicus* (Copepoda): The tidepool copepod, *Tigriopus californicus*, shows exceptional levels of small-scale population genetic heterogeneity associated with the habitat structure of the rocky shoreline, based on studies using mitochondrial markers

(Rawson et al. 2000; Burton et al. 2007). The species may be considered to be a model species for studies of evolutionary divergence and local adaptation (Raisuddin et al. 2007). The rapid rate of evolutionary divergence of mitochondrial genes is thought to contribute to the potential for local adaptation, but may also cause low hybrid fitness by disrupting gene complexes (Burton et al. 2013). The mitochondrial genome has been sequenced (Barreto et al. 2011; Pereira et al. 2016). A genomic SNP linkage map (Foley et al. 2011) and a partial draft genome (https://i5k.nal.usda.gov/Tigriopus_californicus) serve as useful resources for characterizing population genetic diversity and structure. More recently, the capacity of this species to adapt to local condition and stressors has been explored using population genomic and transcriptomic approaches (Lima and Willett 2017; Pereira et al. 2017).

19.4.4. *Acartia tonsa* (Copepoda): The rapid cladogenesis – and perhaps cryptic speciation – of the estuarine copepod, *Acartia tonsa*, has been extensively studied along the Atlantic coastline of the USA using mtDNA marker genes (Caudill and Bucklin 2004; Chen and Hare 2008, 2011). The species has been intensively studied in laboratory culture, partly as food for aquacultured fish (Jepsen et al. 2017) and partly as a model organism for studies of the genetic basis of local adaptation and micro-evolution (Drillet et al. 2008). Responses to environmental stressors have been examined using genomic and transcriptomic approaches (Nilsson et al. 2014; Petkeviciute et al. 2015; Rahlff et al. 2017).

19.4.5. *Euphausia superba* (Euphausiacea): The Antarctic krill, *Euphausia superba* (Fig. 7), is a keystone species of the Southern Ocean pelagic ecosystem, whose high abundance, markedly patchy distribution, and swarming behavior have long been a subject of research (Siegel and Watkins 2016). The population genetic consequences of this exceptional life history have been studied over many decades using varied markers, including allozymes, mitochondrial DNA, and microsatellites. Many studies have revealed similar patterns of genetic diversity, whereby variation within locations far outweighs that between locations, with consistent evidence of lack of large-scale population differentiation (see review by Jarman and Deagle 2016). Two studies using mitochondrial markers found evidence of significant small-scale patchiness: Batta-Lona et al. (2011) hypothesized that genetic differences among samples resulted from advective transport from distinct recruitment centers in the Western Antarctic Peninsula region. Zane et al. (1998) found genetic differentiation between samples collected in the Weddell Sea and South Georgia. Although the statistical significance of these findings has been questioned (see Bortolotto et al. 2011), small-scale patchiness – or genetic "noise" – may be a consequence of the life history of this unique species and/or evidence of local adaptation. Evidence of micro-evolution and local adaptation by Antarctic krill has been shown in genetic and functional analysis of target genes, including thioredoxin (Li et al. 2017), clock genes (Jones and Good 2016), heat shock proteins (Papot et al. 2016), and opsins (Biscontin et al. 2016), among others. Population genomic analysis of Antarctic krill was introduced by Deagle et al. (2015), who examined circum-Antarctic genetic diversity and structure using both RADseq and mitochondrial (ND1 and COI) markers. The large and highly-replicated genome of *E. superba* (47.7 GB, Jeffery 2012) prevented discrimination of allelic variation versus that between copies at separate loci (see above), which was addressed by

analysis of sequence counts at variable nucleotide sites, rather than the derived genotypes. This study confirmed earlier findings of the large-scale panmixia of Antarctic krill populations (Deagle et al. 2015).

19.4.6. *Meganyctiphanes norvegica* (Euphausiacea): The northern krill *Meganyctiphanes norvegica* (Fig. 8) is abundant throughout the North Atlantic and western Mediterranean Sea. The species exhibits clear genetic differentiation among geographic populations based on various mtDNA markers (see review by Patarnello et al. 2010). Consistent evidence of local adaptation of the species, including enzyme activities (Saborowski and Buchholz 2002), is now being analyzed using differential gene expression made possible by a reference transcriptome (Maas and Blanco-Bercial 2016).

19.4.7. *Pleurobrachia bachei* (Ctenophora): A draft genome of the ctenophore *Pleurobrachia bachei* (Fig. 9) revealed the possible preservation of 'ancient molecular toolkits' (Moroz et al. 2014), which are lost in other lineages. The exceptional nature of the genomic architecture of this species can provide new understanding of the genomic basis of their evolutionary success and potential for adaptation. Integrative and comparative analysis of genomic and transcriptomic data of this and another ctenophore species *Mnemiopsis leidyi* demonstrated the phylogenetic position of the phylum as the first metazoan lineage (Ryan et al. 2013; Moroz et al. 2014).

19.4.8. *Spadella cephaloptera* (Chaetognatha): Arrow worms are predatory zooplankton that occupy key positions in pelagic food webs. The phylum comprises many species with cosmopolitan-but-disjunct biogeographical distributions, which has allowed interesting comparisons among species. Population genetic diversity and structure of several chaetognath species have been explored using both mtDNA and microsatellites (Peijnenburg et al. 2004, 2006; Faure and Casanova 2006; Miyamoto et al. 2010; Kulagin et al. 2014). Large-scale studies have also allowed examination of the demographic histories of the species (Peijnenburg et al. 2005). Analysis of the mitochondrial genome of *Spadella cephaloptera* (Fig. 10) yielded evidence of exceptional intraspecific variation (Marlétaz et al. 2017), and resolved the phylogenetic position of the Chaetognatha within Protostome lineages (Papillon et al. 2004).

19.4.9. *Salpa thompsoni* (Tunicata, Thaliacea): The Southern Ocean salp *Salpa thompsoni* (Fig. 11) is a pivotal species in the pelagic ecosystem of Antarctic regions, including the Western Antarctic Peninsula, one of the fastest warming regions of the world oceans. A reference transcriptome for *S. thompsoni* is available, although only 18% of the 216,931 sequences were associated with predicted, hypothetical, or known proteins (Batta-Lona et al. 2017). Another recent study (Jue et al. 2016) produced a preliminary reference genome for the species, identified more than 50% of sequences, and generated both SNP variant and INDEL predictions as a resource for future phylogenetic and population studies. The genome of this species shows evidence of a rapid evolutionary rate – consistent with other Urochordata (Denoeud et al. 2010; Tsagkogeorga et al. 2012). An initial survey of small RNAs revealed the presence of known, conserved miRNAs, novel miRNA genes, and unique piRNA for various

developmental stages (Jue et al. 2016), suggesting possible genomic bases of the successful adaptation of the species to the changing climate of the Southern Ocean.

19.5. Present-day challenges and future opportunities

19.5.1. Additional genomic resources for marine zooplankton species

Pelagic zones represent one of the largest (by volume) habitats on Earth, with highly diverse and ecologically important assemblages of zooplankton, which can serve as early warning indicators of climate change. Genomic resources are needed to facilitate both intra- and interspecies comparative studies of genetic diversity and structure, phylogeography, demographic history, and adaptive evolution. Importantly, marine zooplankton provide a diverse and useful assemblage to move forward novel studies of the genomic basis of adaptation and evolutionary divergence. Yet the exceptional phylogenetic diversity of marine zooplankton exacerbates the challenges of ensuring that reference genomes are available for abundant and ecologically-important species or their close relatives.

Whole-genome sequencing initiatives should cover a wide range of genome sizes to uncover trends in genome evolution and new elements of genome organization. For instance, sequencing of the salp genome revealed novel miRNA genes and unique piRNAs (Jue et al. 2016), while the genome of Pacific sea gooseberry, *Pleurobrachia bachei*, is apparently lacking the canonical miRNA machinery and HOX genes (Moroz et al. 2014).

Stimulating discoveries are anticipated from sequencing the exceptionally large genomes of many crustaceans, including euphausiids, copepods, and amphipods, which may reveal novel regulation of repetitive elements, functional divergence of gene duplication and concomitant novel functions of various gene copies, and correlation between genome size and DNA methylation levels in metazoans (e.g., Lechner et al. 2013). From a practical perspective, even low-coverage genomes will increase the robustness of population genomic approaches by facilitating a diverse range of methods, including in-silico digestion of genome sequences for RAD-seq techniques, higher mapping rates for DNA and RNA-derived sequences, and the development of baits for sequence capture experiments.

Despite their ecological importance in pelagic food webs and their phylogenetic diversity, marine zooplankton have been – and continue to be – largely ignored in the prioritization of species for genomic and transcriptomic analysis. For example, a list of top priority species for reference genome determination from Voolstra et al. (2017) includes only one marine zooplankton species, the mid-water shrimp, *Acanthephyra purpurea*.

19.5.2. Sampling zooplankton in the global ocean

Sampling zooplankton accurately and effectively is a challenge due both to the nature of the pelagic habitat and the frequently immense population sizes of the organisms compared to sampling capacity. It is essential to keep in mind that planktonic organisms most usually occur in patchy distributions, and that some of them are able to avoid the sampling equipment. The origin of these planktonic assemblages or patches has been discussed over many years (e.g., Levin and Segel 1976) and some experimental studies have shown species-specific patterns

(Omori and Hamner 1982). Avoidance behaviors also vary among species, and a number of studies have shown that net size and design can markedly impact avoidance and improve the accuracy of sampling of dense and diverse assemblages (Wiebe 1968; Skjoldal et al. 2013; Wiebe et al. 2013). Novel instrumentation designs are now allowing pairing of net sampling with optical and acoustical technologies to allow adaptive sampling of target species of particular interest and importance.

19.5.3. Species identification

Accurate and precise identification of species is critical for any study, yet for most zooplankton groups this goal is challenging – at best. Morphological identification has been shown to be unreliable for numerous species, including sibling species of the copepods *Pseudocalanus* (Bailey et al. 2015) and *Calanus* (Choquet et al. 2017b). Both transcriptomic and genomic resources are invaluable in allowing the design of rapid and inexpensive protocols for accurate discrimination and identification of sibling and cryptic species of marine zooplankton (e.g., Smolina et al. 2015).

19.5.4. Genomic analysis of small-sized organisms

Zooplankton species are often very small and thus the yield of DNA extractions is limited. This is not an issue for current HTS methods, which usually require a very small amount of DNA (10s ng). The ongoing development of new sequencing platforms and technologies will likely allow longer sequencing reads and thus better genome and transcriptome assemblies. There is a continuing need to ensure that even the tiniest organisms will be amenable to any new developments in sequencing technologies and instrumentation.

19.5.5. Genomic basis of adaptation

Marine environments are experiencing rapid changes in critically-important processes and parameters, including temperature, light penetration, nutrient availability and ocean acidification, among many others. The resultant changes in species physiological condition, ecological functioning, and biogeographical distribution and abundance will inexorably alter pelagic ecosystems in trajectories that are difficult to predict. How species may acclimate and/or adapt to environmental change, and how their interactions within the pelagic food web may be altered, can be examined at many levels. A powerful and important approach lies in examining the underlying genomic mechanisms that facilitate successful adaptation to changing environmental conditions. Although any given species may be uniquely impacted by the physical and biological parameters accompanying shifts in global climate profiles, processes involved in responses to climate change at the molecular level may share common features across species, such as the evolution of gene networks associated with environmental stress responses. Genomic resources are proving instrumental in garnering new insights into organism – environment interactions, including responses to environmental variability associated with climate change. However, we still lack a fundamental understanding of genomic features that afford plasticity and facilitate adaptive responses. These challenges can only be met with comprehensive genomic and transcriptomic resources that will afford comparative analysis to investigate the mechanisms underlying the responses of marine

zooplankton to the changing environmental conditions throughout the global ocean.

19.6. Acknowledgements

This overview results from collaborative efforts with many colleagues and collaborators. We acknowledge and appreciate the contributions of Leocadio Blanco-Bercial (Bermuda Institute of Ocean Sciences), Peter H. Wiebe (Woods Hole Oceanographic Institution, USA), Paola Batta-Lona (CICESE, Mexico), and Nathaniel K. Jue (California State University, USA). Photographic images of living zooplankton species were provided by: L.P. Madin (Woods Hole Oceanographic Institution), J.R. Runge (University of Maine, USA), Slotwinski (University of Tasmania), Uwe Kils (Rutgers University, USA), Dave Wrobel (http://jellieszone.com/), and Peter Parks (Image Quest 3-D). Support was provided by the US National Science Foundation to AB and RJO (PLR-1044982) and to RJO (MCB-1613856); support to IS and MC was provided by Nord University (Norway).

19.7. Table Legends and Tables

Table 1. Holozooplankton species for which genome assemblies and accompanying statistics are publicly available (as of June 2017). Estimated genome sizes are based on assembly statistics unless otherwise noted.

	Tunicates			Ctenophores		Copepods			
	Oikopleura dioica	*Oikopleura dioica*	*Salpa thompsoni*	*Pleurobrachia bachei*	*Mnemiopsis leidyi*	*Eurytemora affinis*	*Oithona nana*	*Tigriopus kingsejongensi*	*Tigriopus californicus*
Assembly Name	Main assem ASM20953v1	Allelic assem ASM20955v1	Salp Genome 1.0	P.bachei_draft_genome_v.1.1	GCA_000226015.1	Eaff_1.0	O. nana v.1.0	NA	TCALIF_v1.0
Estimated Genome Size (MB)	68,46	allelic assembly	602[a]	170	150	616.14[b]	85	298	245[d]
Assembly Size	70 471 451	45 141 193	318 747 957	156 121 975	155 865 547	494 890 867	85 010 107	305 712 242	184 634 130
Predicted protein coding Genes	18 020	18 020	13 186	19 523	16 548	29 783	15 359	12 772	14 536
Coverage	14X	N/A[c]	20X	200X	160X	75X	N/A[c]	65X	N/A[c]
Number of scaffolds	1 260	4 196	478 281	21 979	5 100	6 899	4 626	27 823	2 365
Length of N50 scaffold (bp)	395 387	21 890	934	20 628	187 314	862 645	400 614	159 218	298 012
Number of N50 scaffold (L50)	35	478	79 492	1 646	242	163	60	N/A[c]	180
Number of contigs	5 917	6 678	590 021	38 864	24 884	122 625	7 437	48 368	26 175
Length of N50 contig (bp)	24 932	10 847	636	6 132	11 936	5 738	38 620	17 566	14 799
Number of N50 contig	718	985	136 534	6 078	3 653	19 338	463	N/A[c]	3 352
Length of gaps (bp)	3 938 358	2 655 217	14 945 692	19 276 734	5 525 119	107 316 113	2 943 785	10 474 460	N/A[c]
References	Denoeud et al. (2010)	Denoeud et al. (2010)	Jue et al. (2016)	Moroz et al. (2014)	Ryan et al. (2013)	Eyun et al. (2017)	Madoui et al. (2017)	Kang et al. (2017)	https://i5k.nal.usda.gov/Tigriopus_californicus

[a] genome size estimate independent of assembly (Jue et al. 2016)

[b] genome size estimate independent of assembly (Rasch et al. 2004)

[c] not available/not provided

[d] genome size estimate independent of assembly (Wyngaard and Rasch 2000)

Table 2: Mitochondrial genomes available for marine zooplankton species, with corresponding lengths.

Taxon and Species	Citation	Length (bp)
Copepoda		
Calanus hyperboreus	Kim et al. (2013)	17,910
Calanus sinicus	Minxiao et al. (2011)	>20,460
Paracyclopina nana	Ki et al. (2009)	15,981
Tigriopus californicus	Burton et al. (2007)	14,600
Tigriopus japonicus	Machida et al. (2002)	14,628
Tigriopus sp.	Jung et al. (2006)	14,301
Euphausiacea		
Euphausia pacifica	Shen et al. (2011)	16,898
Euphausia superba	Shen et al. (2010)	>15,498
Ostracoda		
Vargula hilgendorfii	Ogoh & Ohmiya (2004)	15,923
Amphipoda		
Onisimus nanseni	Ki et al. (2010)	14,734
Decapoda		
Acetes chinensis	Kim et al. (2012)	15,740
Cnidaria		
Aurelia aurita	Shao et al. (2006)	16,937
Cassiopea frondosa	Kayal et al. (2011)	15,949
Chrysaora quinquecirrha	Hwang et al. (2014)	16,775
Ctenophora		
Mnemiopsis leidyi	Pett et al. (2011)	10,000
Pleurobrachia bachei	Kohn et al. (2012)	11,016
Chaetognatha		
Sagitta decipiens	Miyamoto et al. (2010)	11,121
Sagitta enflata	Miyamoto et al. (2010)	12,631
Sagitta ferox	Li et al. (2016)	12,153
Sagitta nagae	Miyamoto et al. (2010)	11,459
Paraspadella gotoi	Helfenbein et al. (2004)	11,423
Pterosagitta draco	Wei et al. (2016)	10,426
Spadella cephaloptera	Papillon et al. (2004)	11,905

Table 3. Summary of transcriptomic resources for marine zooplankton species. Transcript and gene numbers are indicated as presented in the original study. Note that different methodologies were employed across these datasets (e.g. Trinity, MIRA_Newbler, Evigene, FPKM filtered, etc) that render cross-comparisons of gene and transcript numbers among species equivocal.

Phylum and Species	**BioProject**	**Contig Total**	**Contig Max Length**	**Contigs Total Length**	**Contigs Annotated**	**Transcripts**	**N50**	**Genes**	**Citation**
Cnidaria									
Alatina alata	PRJNA312373	31 737	32 591	48 508 802	No	31 776	2 545	20 173	Ames et al. (2016)
Rhopilema esculentum	PRJNA318143	148 857	30 742	121 470 903	No	NA	NA	NA	Chongbo and Yunfeng (Dir Sub)
Aurelia aurita	PRJNA252562	252 170	46 960	180 188 094	No	24 264	1 761	10 285	Brekham et al. (2015)
Ctenophora									
Mnemiopsis leidyi	PRJNA344880	140 842	29 348	137 638 938	No	NA	NA	NA	Sanchez Alvarado, Gotting and Ross (Dir Sub)
Arthropoda: Copepoda									
Acartia fossae		100 383	8 174		No		769		Eyun et al. (2017)
Calanus finmarchicus	PRJNA236983	28 954	2 945	10 223 122	No	251 042	354	13 057	Smolina et al. (2014)
Calanus finmarchicus	PRJNA236528	206 012	23 068	205 455 659	Yes		1 418		Lenz et al. (2014)
Calanus finmarchicus	PRJNA231164	241 140	25 048	160 760 719	No				Tarrant et al. (2014)
Calanus glacialis	PRJNA237014	36 880	4 021	15 748 490	No	242 602	471	18 387	Smolina et al. (2014)
Calanus glacialis	PRJNA274584	54 344	7 507	33 214 362	No	16 998	620	16 998	Ramos et al. (2015)
Calanus sinicus		69 751				69 751	1 127	43 417	Yang et al. (2014)
Calanus sinicus			3 923		No	29 458	513		Eyun et al. (2017)
Eucyclops serrulatus	PRJNA231234	51 528	16 342	36 645 141	No				Cattonaro (Dir Sub)
Eurytemora affinis	PRJNA278152	107 445	26 685	142 143 154	No	29 783			Monroe (Dir Sub)
Eurytemora affinis	PRJNA242763	138 088	23 627	143 733 589	Yes				Almada and Tarrant (Dir Sub)
Eurytemora affinis		88 104	26 685						Eyun et al. (2017)
Paracyclopina nana	PRJNA268783	60 687	27 858	95 849 484	Yes	67 179	4 178	12 474	Lee et al. (2015)
Pseudocalanus acuspes	PRJNA296544	207 302	12 713	59 236 626	Yes	69 555	1 348	28 879	De Wit et al. (2016)
Tigriopus kingsejongensis	PRJNA283925	38 250	7 809	36 497 199	Yes				Lee (Dir Sub)
Tigriopus kingsejongensis			23 942	28 850 726		40 172	1 093	12 772	Kang et al. (2017)
Tigriopus californicus	PRJNA263967	12 067	13 452	14 966 851	No				Barreto et al. (2011)
Tigriopus californicus	PRJNA263967	12 075	13 452	14 902 878	No				Barreto et al. (2011)
Tigriopus californicus		106 317	27 644	NA	Yes	106 317	2 837	12 573	Periera et al. (2016)
Tigriopus californicus		60 840	8 614				1 510		Eyun et al. (2017)
Tigriopus japonicus	PRJNA274317	54 758	23 769	82 981 758	Yes		3 565		Kim et al. (2015)
Arthropoda: Euphausiacea									
Euphausia superba			11 127		Yes	15 347	520	7 942	Meyer et al. (2015)
Euphausia superba		22 177	8 515		Yes			5 563	Clark et al. (2011)
Euphausia superba		133 962		129 183 922	Yes		1 294	27 928	Sales et al (2017)
Euphasia crystallorophias		42 632					8 341		Toullec et al. (2013)
Meganyctiphanes norvegica	PRJNA324094	405 497	26 644	222 530 071	No	NA	NA	NA	Maas and Blanco Bercial (Dir Sub)
Arthropoda: Amphipoda									
Talitrus saltator	PRJNA297565	156 706	22 032	151 674 147	Yes		968		O'Grady et al. (2016)
Arthropoda: Mysidacea									
Neomysis awatschensis	PRJNA287057	22 141	10 398	14 999 154	Yes	22 141	801		Kim et al. (2016)
Mollusca: Pteropoda									
Clio pyramidata	PRJNA231010	45 735				45 735	852	30 800	Maas et al. (2015)
Clione limacina	PRJNA314884	477 401	30 190	258 267 445	Yes	300 994	816	181 879	Thabet et al. (2017)
Limacina antarctica	PRJNA295792	81 226	7 935	59791880	No	402 273	500	81 229	Johnson and Hoffman (2016)
Limacina helicina	PRJNA386290	53 121	12 358	31 790 000	Yes		796		Koh et al. (2015)
Urochordata: Tunicata									
Oikopleura dioica	PRJNA269316	54 949	23 096	66 526 340	No				Wang et al. (2015)
Oikopleura dioica	PRJNA269317	86 898		70 800 000		57 962	1 806	16 423	Wang et al. (2015)
Salpa thompsoni	PRJNA279245	217 849	30 785	151 741 986	No	216 931	1 163	26 413	Jue et al. (2016); Batta Lona et al. (2017)

19.8. Figure Legends

Figure 1. Distribution of estimated genome sizes in representative holozooplankton phyla. Black dots indicate sequenced genomes. Genome size estimations are from Gregory (2017), Jeffery et al. (2017), Leinaas et al. (2016), Ryan et al. (2014), Moroz et al. (2014), and Madoui et al. (2017).

Figure 2: Comparison of the mitochondrial gene order between *Calanus sinicus* and *C. hyperboreus*. Only the 13 protein-coding genes are represented. Rectilinear shapes show genes for which the order is conserved between the two species; red lines indicate genes with the same sequence but in reverse order between the species.

Figure 3. *Calanus finmarchicus* (Copepod) http://umaine.edu/jrunge/files/2013/12/CV_1_for-publication.jpg (Photo J.R. Runge, University of Maine)

Figure 4. Circulation patterns and bathymetry of the North Atlantic Ocean basin, providing the foundation of the three-gyre hypothesis for basin-scale dispersal of the copepod *Calanus finmarchicus*. Figure from Wiebe et al. (2009).

Figure 5. *Centropages typicus* (Copepod) https://alchetron.com/Centropages-2143715-W (Photo Slotwinski, University of Tasmania)

Figure 6. Hypothesized models of gene flow and population connectivity of the copepod *Centropages typicus*. The full migration model (upper right in diagram) showed the highest likelihood among the considered models based on Bayesian analysis. Figure from Blanco-Bercial and Bucklin (2016).

Figure 7. *Euphausia superba* (Euphausiid) http://www.ecoscope.com/krill/krill4/index.htm (Photo Uwe Kils, Rutgers University, USA)

Figure 8. *Meganyctiphanes norvegica* (Euphausiid) https://en.wikipedia.org/wiki/Northern_krill#/media/File:Meganyctiphanes_norvegica.jpg (Photo Uwe Kils, Rutgers University, USA)

Figure 9. *Pleurobrachia bachei* (Ctenophora) http://jellieszone.com/ctenophores/pleurobrachia/ (Photo Dave Wrobel)

Figure 10. *Spadella cephaloptera* (Chaetognatha) http://australianmuseum.net.au/image/Arrow-worm-Chaetognaths (Photo Peter Parks, Image Quest 3-D)

Figure 11. *Salpa thompsoni* (Tunicata, Thaliacea) http://www.whoi.edu/cms/images/oceanus/2005/6/v44n1-briefs2-3en_10823.jpg (Photo L.P. Madin, Woods Hole Oceanographic Institution)

19.9 Literature Cited

Aarbakke ONS, Bucklin A, Halsband C, Norrbin F (2011) Discovery of *Pseudocalanus moultoni* (Frost, 1989) in Northeast Atlantic waters based on mitochondrial COI sequence variation. J Plankton Res 33:1487–1495. doi: 10.1093/plankt/fbr057

Aarbakke ONS, Bucklin A, Halsband C, Norrbin F (2014) Comparative phylogeography and demographic history of five sibling species of *Pseudocalanus* (Copepoda: Calanoida) in the North Atlantic Ocean. J Exp Mar Bio Ecol 461:479–488. doi: 10.1016/j.jembe.2014.10.006

Abad D, Albaina A, Aguirre M, et al (2016) Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Mar Biol. doi: 10.1007/s00227-016-2920-0

Alberto F, Raimondi PT, Reed DC, et al (2011) Isolation by oceanographic distance explains genetic structure for *Macrocystis pyrifera* in the Santa Barbara Channel. Mol Ecol 20:2543–2554. doi: 10.1111/j.1365-294X.2011.05117.x

Alfsnes K, Leinaas HP, Hessen DO (2017) Genome size in arthropods; different roles of phylogeny, habitat and life history in insects and crustaceans. Ecol Evol 1–9. doi: 10.1002/ece3.3163

Ames CL, Ryan JF, Bely AE, et al (2016) A new transcriptome and transcriptome profiling of adult and larval tissue in the box jellyfish *Alatina alata*: an emerging model for studying venom, vision and sex. BMC Genomics 17:650. doi: 10.1186/s12864-016-2944-3

Andrews KR, Norton EL, Fernandez-Silva I, et al (2014) Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. Mol Ecol 23:5462–5479. doi: 10.1111/mec.12950

Avise JC (2009) Phylogeography: Retrospect and prospect. J Biogeogr 36:3–15. doi: 10.1111/j.1365-2699.2008.02032.x

Avise JC, Bowen BW, Ayala FJ (2016) In the light of evolution X: Comparative phylogeography. Proc Natl Acad Sci U S A 113:7957–7961. doi: 10.1073/pnas.1604338113

Bailey J, Rynearson T, Durbin EG (2015) Species composition and abundance of copepods in the morphologically cryptic genus *Pseudocalanus* in the Bering Sea. Deep Sea Res Part II Top Stud Oceanogr in press. doi: 10.1016/j.dsr2.2015.04.017

Baird NA, Etter PD, Atwood TS, et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS One 3:1–7. doi: 10.1371/journal.pone.0003376

Barreto FS, Moy GW, Burton RS (2011) Interpopulation patterns of divergence and selection across the transcriptome of the copepod *Tigriopus californicus*. Mol Ecol 20:560–572. doi: 10.1111/j.1365-294X.2010.04963.x

Batta-Lona PG, Bucklin A, Wiebe PH, et al (2011) Population genetic variation of the Southern Ocean krill, *Euphausia superba*, in the Western Antarctic Peninsula region based on mitochondrial single nucleotide polymorphisms (SNPs). Deep Sea Res Part II Top Stud Oceanogr 58:1652–1661. doi: 10.1016/j.dsr2.2010.11.017

Batta-Lona PG, Maas AE, O'Neill RJ, et al (2017) Transcriptomic profiles of spring and summer populations of the Southern Ocean salp, *Salpa thompsoni*, in the Western Antarctic Peninsula region. Polar Biol 40:1261–1276. doi: 10.1007/s00300-016-2051-6

Beaugrand G (2017) Plankton biodiversity and biogeography. In: Castellani C, Edwards M (eds) Marine Plankton: A Practical Guide to Ecology, Methodology, and Taxonomy. Oxford University Press, pp 12–23

Beerli P (2012) Migrate Documentation version 3.2.1. Florida State University, Tallahasee. Florida

Bierne N, Bonhomme F, Arnaud-Haond AS (2016) Dedicated population genomics for the silent world: The specific questions of marine population genetics. Curr Zool 62:545–550. doi: 10.1093/cz/zow107

Bierne N, Roze D, Welch JJ (2013) Pervasive selection or is it …? why are FST outliers sometimes so frequent? Mol Ecol 33:2061–2064. doi: 10.1111/mec.12241

Bik HM, Porazinska DL, Creer S, et al (2012) Sequencing our way towards understanding global eukaryotic biodiversity. Trends Ecol Evol 27:233–243. doi: 10.1016/j.tree.2011.11.010

Biscontin A, Frigato E, Sales G, et al (2016) The opsin repertoire of the Antarctic krill *Euphausia superba*. Mar Genomics 29:61–68. doi: 10.1016/j.margen.2016.04.010

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. Annu Rev Entomol 46:441–469. doi: 10.1146/annurev.ento.46.1.441

Blanco-Bercial L, Álvarez-Marqués F, Bucklin A (2011a) Comparative phylogeography and connectivity of sibling species of the marine copepod *Clausocalanus* (Calanoida). J Exp Mar Bio Ecol 404:108–115. doi: 10.1016/j.jembe.2011.05.011

Blanco-Bercial L, Álvarez-Marqués F, Bucklin A (2011b) Comparative phylogeography and connectivity of sibling species of the marine copepod *Clausocalanus* (Calanoida). J Exp Mar Bio Ecol 404:108–115. doi: 10.1016/j.jembe.2011.05.011

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

Blanco-Bercial L, Cornils A, Copley N, Bucklin A (2014) DNA Barcoding of marine copepods: assessment of analytical approaches to species identification. PLoS Curr 1–40. doi: 10.1371/currents.tol.cdf8b74881f87e3b01d56b43791626d2

Bolte S, Fuentes V, Haslob H, et al (2013) Population genetics of the invasive ctenophore *Mnemiopsis leidyi* in Europe reveal source-sink dynamics and secondary dispersal to the Mediterranean Sea. Mar Ecol Prog Ser 485:25–36. doi: 10.3354/meps10321

Bortolotto E, Bucklin A, Mezzavilla M, et al (2011) Gone with the currents: lack of genetic differentiation at the circum-continental scale in the Antarctic krill *Euphausia superba*. BMC Genet 12:32. doi: 10.1186/1471-2156-12-32

Brekhman V, Malik A, Haas B, et al (2015) Transcriptome profiling of the dynamic life cycle of the scypohozoan jellyfish *Aurelia aurita*. BMC Genomics 16:74. doi: 10.1186/s12864-015-1320-z

Bron JE, Frisch D, Goetze E, et al (2011) Observing copepods through a genomic lens. Front Zool 8:22. doi: 10.1186/1742-9994-8-22

Bucklin A, Hopcroft RR, Kosobokova KN, et al (2010a) DNA barcoding of Arctic Ocean holozooplankton for species identification and recognition. Deep Res Part II Top Stud Oceanogr 57:40–48. doi: 10.1016/j.dsr2.2009.08.005

Bucklin A, Kaartvedt S (2000) Population genetics of drifting (*Calanus* spp.) and resident (*Acartia clausi*) plankton in Norwegian fjords. J Plankton Res 22:1237–1251. doi: 10.1093/plankt/22.7.1237

Bucklin A, Lindeque PK, Rodriguez-Ezpeleta N, et al (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. J Plankton Res 38:393–400. doi: 10.1093/plankt/fbw023

Bucklin A, Ortman BD, Jennings RM, et al (2010b) A "Rosetta Stone" for metazoan zooplankton: DNA barcode analysis of species diversity of the Sargasso Sea (Northwest Atlantic Ocean). Deep Res Part II Top Stud Oceanogr 57:2234–2247. doi: 10.1016/j.dsr2.2010.09.025

Bucklin A, Smolenack SB, Bentley AM, Wiebe PH (1997) Gene flow patterns of the euphausiid, *Meganyctiphanes norvegica*, in the NW Atlantic based on mtDNA sequences for cytochrome b and cytochrome oxidase I. J Plankton Res 19:1763–1781. doi: 10.1093/plankt/19.11.1763

Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. Ann Rev Mar Sci 3:471–508. doi: 10.1146/annurev-marine-120308-080950

Bucklin A, Sundt RC, Dahle G (1996) The population genetics of *Calanus finmarchicus* in the North Atlantic. Ophelia 44:29–45.

Bucklin A, Wiebe PH, Smolenack SB, et al (2007) DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). J Plankton Res 29:483–493. doi: 10.1093/plankt/fbm031

Burton RS, Byrne RJ, Rawson PD (2007) Three divergent mitochondrial genomes from California populations of the copepod *Tigriopus californicus*. Gene 403:53–59. doi: 10.1016/j.gene.2007.07.026

Burton RS, Pereira RJ, Barreto FS (2013) Cytonuclear genomic interactions and hybrid breakdown. Annu Rev Ecol Evol Syst 44:281–302. doi: 10.1146/annurev-ecolsys-110512-135758

Bybee SM, Bracken-Grissom HD, Hermansen RA, et al (2011) Directed next generation sequencing for phylogenetics: An example using Decapoda (Crustacea). Zool Anz 250:497–506. doi: 10.1016/j.jcz.2011.05.010

Carlotti F, Bonnet D, Halsband-Lenk C (2007) Development and growth rates of *Centropages typicus*. Prog Oceanogr 72:164–194. doi: 10.1016/j.pocean.2007.01.011

Castellani C, Lindley AJ, Wootton M, et al (2012) Morphological and genetic variation in the North Atlantic copepod, *Centropages typicus*. J Mar Biol Assoc United Kingdom 92:99–106. doi: 10.1017/S0025315411000932

Caudill CC, Bucklin A (2004) Molecular phylogeography and evolutionary history of the estuarine copepod, *Acartia tonsa*, on the Northwest Atlantic coast. Hydrobiologia 511:91–102. doi: 10.1023/B:HYDR.0000014032.05680.9d

Chen G, Hare MP (2008) Cryptic ecological diversification of a planktonic estuarine copepod, A*cartia tonsa*. Mol Ecol 17:1451–1468. doi: 10.1111/j.1365-294X.2007.03657.x

Chen G, Hare MP (2011) Cryptic diversity and comparative phylogeography of the estuarine copepod *Acartia tonsa* on the US Atlantic coast. Mol Ecol 20:2425–2441. doi: 10.1111/j.1365-294X.2011.05079.x

Choquet M, Alves Monteiro HJ, Bengtsson-Palme J, Hoarau G (2017a) The complete mitochondrial genome of the copepod *Calanus glacialis*. Mitochondrial DNA Part B 2:506–507. doi: 10.1080/23802359.2017.1361357

Choquet M, Smolina I, Soreide JE, Hoarau G (2017b) New insight on the population structure of *Calanus finmarchicus* in the North Atlantic using next-generation sequencing technologies. In: 13th International Conference on Copepoda. Los Angeles, CA, USA,

Chust G, Villarino E, Chenuil A, et al (2016) Dispersal similarly shapes both population genetics and community patterns in the marine realm. Sci Rep 6:28730. doi: 10.1038/srep28730

Clark MS, Thorne MAS, Toullec JY, et al (2011) Antarctic krill 454 pyrosequencing reveals chaperone and stress transcriptome. PLoS One 6:1–17. doi: 10.1371/journal.pone.0015919

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

Cowen RK, Sponaugle S (2009) Larval Dispersal and Marine Population Connectivity. Annu Rev Mar Sci 1:443–66. doi: 10.1146/annurev.marine.010908.163757

Crawford DL, Oleksiak MF (2016) Ecological population genomics in the marine environment. Brief Funct Genomics 15:342–351. doi: 10.1093/bfgp/elw008

Cristescu ME (2015) Genetic reconstructions of invasion history. Mol Ecol 24:2212–2225. doi: 10.1111/mec.13117

Davey JL, Blaxter MW (2010) RADseq: Next-generation population genetics. Brief Funct Genomics 9:416–423. doi: 10.1093/bfgp/elq031

Dawson MN, Cieciel K, Decker MB, et al (2015) Population-level perspectives on global change: genetic and demographic analyses indicate various scales, timing, and causes of scyphozoan jellyfish blooms. Biol Invasions 17:851–867. doi: 10.1007/s10530-014-0732-z

Deagle BE, Faux C, Kawaguchi S, et al (2015) Antarctic krill population genomics: Apparent panmixia, but genome complexity and large population size muddy the water. Mol Ecol 24:4943–4959. doi: 10.1111/mec.13370

De Pittà C, Biscontin A, Albiero A, et al (2013) The Antarctic Krill *Euphausia superba* Shows Diurnal Cycles of Transcription under Natural Conditions. PLoS One. doi: 10.1371/journal.pone.0068652

de Vargas C, Audic S, Henry N, et al (2015) Eukaryotic plankton diversity in the sunlit ocean. Science 348:1261605. doi: 10.1126/science.1261605

De Wit P, Dupont S, Thor P (2016) Selection on oxidative phosphorylation and ribosomal structure as a multigenerational response to ocean acidification in the common copepod *Pseudocalanus acuspes*. Evol Appl 9:1112–1123. doi: 10.1111/eva.12335

De Wit P, Pespeni MH, Palumbi SR (2015) SNP genotyping and population genomics from expressed sequences - Current advances and future possibilities. Mol Ecol 24:2310–2323. doi: 10.1111/mec.13165

Denoeud F, Henriet S, Mungpakdee S (2010) Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. Science (80-) 1381–1386.

deYoung B, Barange M, Beaugrand G, et al (2008) Regime shifts in marine ecosystems: detection, prediction and management. Trends Ecol Evol 23:402–409. doi: 10.1016/j.tree.2008.03.008

Drillet G, Goetze E, Jepsen PM, et al (2008) Strain-specific vital rates in four *Acartia tonsa* cultures, I: Strain origin, genetic differentiation and egg survivorship. Aquaculture 280:109–116. doi: 10.1016/j.aquaculture.2008.04.005

Dufresne F, Jeffery N (2011) A guided tour of large genome size in animals: What we know and where we are heading. Chromosom Res 19:925–938. doi: 10.1007/s10577-011-9248-x

Edmands S (2001) Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. Mol Ecol 10:1743–1750. doi: 10.1046/j.0962-1083.2001.01306.x

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

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. Trends Ecol Evol 29:51–63. doi: 10.1016/j.tree.2013.09.008

Elliott TA, Gregory TR (2015) What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. Philos Trans R Soc London Ser B-Biological Sci 370:20140331. doi: 10.1098/rstb.2014.0331

Elshire RJ, Glaubitz JC, Sun Q, et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:1–10. doi: 10.1371/journal.pone.0019379

Escribano R, McLaren IA, Breteler WCMK (1992) Innate and acquired variation of nuclear DNA contents of marine copepods. Genome 35:602–610. doi: 10.1139/g92-090

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479–491. doi: 10.1007/s00424-009-0730-7

Eyun S, Young Soh H, Posavi M, et al (2017) Evolutionary History of Chemosensory-Related Gene Families Across the Arthropoda. Mol Biol Evol 34:1838–1862. doi: 10.1093/molbev/msx147

Falk-Petersen S, Pavlov V, Timofeev S, Sargent J (2007) Climate variability and possible effects on Arctic food chains: the role of *Calanus*. In: Arctic Apline ecosystems and People in a Changing Environment. Springer Berlin Heidelberg, pp 147–166

Faure E, Casanova JP (2006) Comparison of chaetognath mitochondrial genomes and phylogenetical implications. Mitochondrion 6:258–262. doi: 10.1016/j.mito.2006.07.004

Foley BR, Rose CG, Rundle DE, et al (2011) A gene-based SNP resource and linkage map for the copepod *Tigriopus californicus*. BMC Genomics 12:568. doi: 10.1186/1471-2164-12-568

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180:977–993. doi: 10.1534/genetics.108.092221

Francisco SM, Robalo JI, Levy A, Almada VC (2014) In Search of Phylogeographic Patterns in the Northeastern Atlantic and Adjacent Seas. In: Evolutionary Biology: Genome Evolution, Speciation, Coevolutions and Origin of Life. Springer Publishing, pp 323–338

Gagnaire PA, Broquet T, Aurelle D, et al (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. Evol Appl 8:769–786. doi: 10.1111/eva.12288

Gagnaire PA, Gaggiotti OE (2016) Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. Curr Zool 62:603–616. doi: 10.1093/cz/zow088

Galindo HM, Pfeiffer-Herbert AS, McManus MA, et al (2010) Seascape genetics along a steep cline: Using genetic patterns to test predictions of marine larval dispersal. Mol Ecol 19:3692–3707. doi: 10.1111/j.1365-294X.2010.04694.x

Gayral P, Melo-Ferreira J, Glémin S, et al (2013) Reference-Free Population Genomics from Next-Generation Transcriptome Data and the Vertebrate-Invertebrate Gap. PLoS Genet. doi: 10.1371/journal.pgen.1003457

Genome 10K Community of Scientists (2009) Genome 10K: A proposal to obtain whole-genome sequence for 10000 vertebrate species. J Hered 100:659–674. doi: 10.1093/jhered/esp086

GIGA Community of Scientists (2014) The Global Invertebrate Genomics Alliance (GIGA): Developing Community Resources to Study Diverse Invertebrate Genomes. J Hered 105:1–18. doi: 10.1093/jhered/est084

Goetze E (2005) Global population genetic structure and biogeography of the oceanic copepods *Eucalanus hyalinus* and *E. spinifer*. Evolution 59:2378–2398. doi: 10.1554/05-077.1

Goetze E, Andrews KR, Peijnenburg KTCA, et al (2015) Temporal stability of genetic structure in a mesopelagic copepod. PLoS One 10:1–16. doi: 10.1371/journal.pone.0136087

Goetze E, Ohman MD (2010) Integrated molecular and morphological biogeography of the calanoid copepod family Eucalanidae. Deep Res Part II Top Stud Oceanogr 57:2110–2129. doi: 10.1016/j.dsr2.2010.09.014

Gregory TR (2017) Animal Genome Size Databse. http://www.genomesize.com. Accessed 20 Jun 2017

Gregory TR, Hebert PDN (1999) The modulation of DNA content: Proximate causes and ultimate consequences. Genome Res 9:317–324. doi: 10.1101/gr.9.4.317

Gregory TR, Hebert PDN, Kolasa J (2000) Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. Heredity (Edinb) 84:201–208. doi: 10.1046/j.1365-2540.2000.00661.x

Hahn C, Bachmann L, Chevreux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. Nucleic Acids Res. doi: 10.1093/nar/gkt371

Hansen MM, Hemmer-Hansen J (2007) Landscape genetics goes to sea. J Biol 6:6. doi: 10.1186/jbiol59

Havird JC, Santos SR (2016) Here we are, but where do we go? A systematic review of crustacean transcriptomic studies from 2014-2015. Integr Comp Biol 56:1055–1066. doi: 10.1093/icb/icw061

Head EJH, Harris LR, Yashayaev I (2003) Distributions of *Calanus* spp. and other mesozooplankton in the Labrador Sea in relation to hydrography in spring and summer (1995-2000). Prog Oceanogr 59:1–30. doi: 10.1016/S0079-6611(03)00111-3

Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: The age of genomics. Annu Rev Ecol Evol Syst 37:67–93. doi: 10.1146/annurev.ecolsys.37.091305.110132

Helfenbein KG, Fourcade HM, Vanjani RG, Boore JL (2004) The mitochondrial genome of *Paraspadella gotoi* is highly reduced and reveals that chaetognaths are a sister group to protostomes. Proc Natl Acad Sci U S A 101:10639–10643. doi: 10.1073/pnas.0400941101

Hellberg ME (2009) Gene Flow and Isolation among Populations of Marine Animals. Annu Rev Ecol Evol Syst 40:291–310. doi: DOI 10.1146/annurev.ecolsys.110308.120223

Helyar SJ, Hemmer-Hansen J, Bekkevold D, et al (2011) Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. Mol Ecol Resour 11:123–136. doi: 10.1111/j.1755-0998.2010.02943.x

Hemmer-Hansen J, Overgaard N, Hemmer-Hansen J, et al (2014) Population genomics of marine fishes: next generation prospects and challenges Population Genomics of Marine Fishes: Next-Generation Prospects and Challenges. Biol Bull 227:117–132. doi: 10.1086/BBLv227n2p117

Hessen DO, Persson J (2009) Genome size as a determinant of growth and life-history traits in crustaceans. Biol J Linn Soc 98:393–399. doi: 10.1111/j.1095-8312.2009.01285.x

Hirai J, Tsuda A (2015) Metagenetic community analysis of epipelagic planktonic copepods in the tropical and subtropical pacific. Mar Ecol Prog Ser 534:65–78. doi: 10.3354/meps11404

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

Hodges E, Xuan Z, Balija V, et al (2007) Genome-wide in situ exon capture for selective resequencing. Nat Genet 39:1522–7. doi: 10.1038/ng.2007.42

Hwang D-S, Park E, Won Y-J, et al (2014) Complete mitochondrial genome of the jellyfish, *Chrysaora quinquecirrha* (Cnidaria, Scyphozoa). Mitochondrial DNA 25:25–26. doi: 10.3109/19401736.2013.775272

i5K Consortium (2013) The i5K initiative: Advancing arthropod genomics for knowledge, human health, agriculture, and the environment. J Hered 104:595–600. doi: 10.1093/jhered/est050

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

Jarman SN, Deagle BE (2016) Genetics of Antarctic krill. In: Siegel V (ed) Biology and Ecology of Antarctic Krill. Advances in Polar Ecology, pp 247–277

Jeffery NW (2012) The first genome size estimates for six species of krill (Malacostraca, Euphausiidae): Large genomes at the north and south poles. Polar Biol 35:959–962. doi: 10.1007/s00300-011-1137-4

Jeffery NW, Ellis EA, Oakley TH, Gregory TR (2017) The Genome Sizes of Ostracod Crustaceans Correlate with Body Size and Evolutionary History, but not Environment. J Hered 108:701–706. doi: 10.1093/jhered/esx055

Jeffrey NW (2015) Genome Size Diversity and Evolution in the Crustacea. University of Guelph, Canada

Jepsen PM, Bjørbæk NS, Rayner TA, et al (2017) Recommended feeding regime and light climate in live feed cultures of the calanoid copepod *Acartia tonsa* Dana. Aquac Int 25:635–654. doi: 10.1007/s10499-016-0063-4

Johnson KM, Hofmann GE (2016) A transcriptome resource for the Antarctic pteropod *Limacina helicina antarctica*. Mar Genomics 28:25–28. doi: 10.1016/j.margen.2016.04.002

Jones MR, Good JM (2016) Targeted capture in evolutionary and ecological genomics. Mol Ecol 25:185–202. doi: 10.1111/mec.13304

Jue NK, Batta-Lona PG, Trusiak S, et al (2016) Rapid evolutionary rates and unique genomic signatures discovered in the first reference genome for the southern ocean salp, *salpa thompsoni* (Urochordata, Thaliacea). Genome Biol Evol 8:3171–3186. doi: 10.1093/gbe/evw215

Jung SO, Lee YM, Park TJ, et al (2006) The complete mitochondrial genome of the intertidal copepod *Tigriopus* sp. (Copepoda, Harpactidae) from Korea and phylogenetic considerations. J Exp Mar Bio Ecol 333:251–262. doi: 10.1016/j.jembe.2005.12.047

Kang S, Ahn D, Lee JH, et al (2017) The genome of the Antarctic-endemic copepod, *Tigriopus kingsejongensis*. Gigascience 1–9. doi: 10.1093/gigascience/giw010

Ki JS, Hop H, Kim SJ, et al (2010) Complete mitochondrial genome sequence of the Arctic gammarid, *Onisimus nanseni* (Crustacea; Amphipoda): Novel gene structures and unusual control region features. Comp Biochem Physiol - Part D Genomics Proteomics 5:105–115. doi: 10.1016/j.cbd.2010.02.002

Ki JS, Lee KW, Park HG, et al (2009) Phylogeography of the copepod *Tigriopus japonicus* along the Northwest Pacific rim. J Plankton Res 31:209–221. doi: 10.1093/plankt/fbn100

Kim HS, Lee BY, Won EJ, et al (2015) Identification of xenobiotic biodegradation and metabolism-related genes in the copepod *Tigriopus japonicus* whole transcriptome analysis. Mar Genomics 24:207–208. doi: 10.1016/j.margen.2015.05.011

Kim S, Kim J, Choi H-G, et al (2012) Complete mitochondrial genome of the northern mauxia shrimp *Acetes chinensis* (Decapoda, Dendrobranchiata, Sergestoidae). Mitochondrial DNA 23:28–30. doi: 10.3109/19401736.2011.643878

Kim S, Lim BJ, Min GS, Choi HG (2013) The complete mitochondrial genome of Arctic *Calanus hyperboreus* (Copepoda, Calanoida) reveals characteristic patterns in calanoid mitochondrial genome. Gene 520:64–72. doi: 10.1016/j.gene.2012.09.059

Knowles LL (2009) Statistical Phylogeography. Annu Rev Ecol Evol Syst 40:593–612. doi: 10.1146/annurev.ecolsys.38.091206.095702

Koh HY, Lee JH, Han SJ, et al (2015) A transcriptomic analysis of the response of the arctic pteropod *Limacina helicina* to carbon dioxide-driven seawater acidification. Polar Biol 38:1727–1740. doi: 10.1007/s00300-015-1738-4

Kohn AB, Citarella MR, Kocot KM, et al (2012) Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, *Pleurobrachia bachei*. Mol Phylogenet Evol 63:203–207. doi: 10.1016/j.ympev.2011.12.009

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

Kool JT, Moilanen A, Treml EA (2013) Population connectivity: Recent advances and new perspectives. Landsc Ecol 28:165–185. doi: 10.1007/s10980-012-9819-z

Kulagin DN, Stupnikova AN, Neretina T V., Mugue NS (2014) Spatial genetic heterogeneity of the cosmopolitan chaetognath *Eukrohnia hamata* (Möbius, 1875) revealed by mitochondrial DNA. Hydrobiologia 721:197–207. doi: 10.1007/s10750-013-1661-z

Kuriyama M, Nishida S (2006) Species diversity and niche-partitioning in the pelagic copepods of the family Scolecitrichidae (Calanoida). Crustaceanna 79:293–317.

Laakmann S, Auel H, Kochzius M (2012) Evolution in the deep sea: Biological traits, ecology and phylogenetics of pelagic copepods. Mol Phylogenet Evol 65:535–546. doi: 10.1016/j.ympev.2012.07.007

Lauritano C, Procaccini G, Ianora A (2012) Gene expression patterns and stress response in marine copepods. Mar Environ Res 76:22–31. doi: 10.1016/j.marenvres.2011.09.015

Lechner M, Marz M, Ihling C, et al (2013) The correlation of genome size and DNA methylation rate in metazoans. Theory Biosci 132:47–60. doi: 10.1007/s12064-012-0167-y

Lee BY, Kim HS, Choi BS, et al (2015) RNA-seq based whole transcriptome analysis of the cyclopoid copepod *Paracyclopina nana* focusing on xenobiotics metabolism. Comp Biochem Physiol - Part D Genomics Proteomics 15:12–19. doi: 10.1016/j.cbd.2015.04.002

Lee CE (2000) Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". Evolution 54:2014–2027. doi: 10.1111/j.0014-3820.2000.tb01245.x

Lee CE (2016) Evolutionary mechanisms of habitat invasions, using the copepod *Eurytemora affinis* as a model system. Evol Appl 9:248–270. doi: 10.1111/eva.12334

Leinaas HP, Jalal M, Gabrielsen TM, Hessen DO (2016) Inter- and intraspecific variation in body- and genome size in calanoid copepods from temperate and arctic waters. Ecol Evol 6:5585–5595. doi: 10.1002/ece3.2302

Lenz PH, Roncalli V, Hassett RP, et al (2014) De novo assembly of a transcriptome for *Calanus finmarchicus* (crustacea, copepoda) - The dominant zooplankter of the North Atlantic Ocean. PLoS One. doi: 10.1371/journal.pone.0088589

Leray M, Knowlton N (2016) Censusing marine eukaryotic diversity in the twenty-first century. Philos Trans R Soc London Ser B, Biol Sci R Soc London Ser B-Biological Sci 371:1–9. doi: 10.1098/rstb.2015.0331

Levasseur A, Orlando L, Bailly X, et al (2007) Conceptual bases for quantifying the role of the environment on gene evolution: The participation of positive selection and neutral evolution. Biol Rev 82:551–572. doi: 10.1111/j.1469-185X.2007.00024.x

Levin SA, Segel LA (1976) Hypothesis for origin of planktonic patchiness. Nature 259:659–659. doi: 10.1038/259659a0

Li F, Ma L, Zhang H, et al (2017) A thioredoxin from Antarctic microcrustacean (*Euphausia superba*): Cloning and functional characterization. Fish Shellfish Immunol 63:376–383. doi: 10.1016/j.fsi.2017.02.035

Li P, Yang M, Ni S, et al (2016) Complete mitochondrial genome sequence of the pelagic chaetognath, *Sagitta ferox*. Mitochondrial DNA 1736:1–2. doi: 10.3109/19401736.2015.1106508

Lima TG, Willett CS (2017) Locally adapted populations of a copepod can evolve different gene expression patterns under the same environmental pressures. Ecol Evol 7:4312–4325. doi: 10.1002/ece3.3016

Lindeque PK, Parry HE, Harmer RA, et al (2013) Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PLoS One 8:1–14. doi: 10.1371/journal.pone.0081327

Longhurst AR (2007) Ecological Geography of the Sea. Elsevier Inc.

Luikart G, England PR, Tallmon D, et al (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981–994. doi: 10.1038/nrg1226

Maas AE, Blanco-Bercial L (2016) *Meganyctiphanes norvegica* Transcriptome and Gene expression, GenBank Direct Accession No. PRJNA324094 ID:324094. https://www.ncbi.nlm.nih.gov/bioproject/?term=Meganyctiphanes norvegica.

Maas AE, Lawson GL, Tarrant AM (2015) Transcriptome-wide analysis of the response of the thecosome pteropod *Clio pyramidata* to short-term CO_2 exposure. Comp Biochem Physiol - Part D Genomics Proteomics 16:1–9. doi: 10.1016/j.cbd.2015.06.002

Machida RJ, Miya MU, Nishida M, Nishida S (2002) Complete mitochondrial DNA sequence of *Tigriopus japonicus* (Crustacea: Copepoda). Mar Biotechnol 4:406–417. doi: 10.1007/s10126-002-0033-x

Madoui M-A, Poulain J, Sugier K, et al (2017) New insights into global biogeography, population structure and natural selection from the genome of the epipelagic copepod *Oithona*. Mol Ecol 38:42–49. doi: 10.1111/mec.14214

Maricic T, Whitten M, Pääbo S (2010) Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One 5:9–13. doi: 10.1371/journal.pone.0014004

Marlétaz F, Parco Y., Shenglin L, Peijnenburg KTCA (2017) Extreme Mitogenomic Variation in Natural Populations of Chaetognaths. Genome Biol Evol 9:1–21. doi: 10.1093/gbe/evx090

McGovern TM, Keever CC, Saski CA, et al (2010) Divergence genetics analysis reveals historical population genetic processes leading to contrasting phylogeographic patterns in co-distributed species. Mol Ecol 19:5043–5060. doi: 10.1111/j.1365-294X.2010.04854.x

McLaren I a, Laberge E, Corkett CJ, Sevigny J-M (1989) Life cycles of four species of *Pseudocalanus* in Nova Scotia. Can J Zool 67:552–558. doi: doi:10.1139/z89-078

McLaren IA, Sevigny J-M, Corkett CJ (1988) Body sizes, development rates, and genome sizes among *Calanus* species. Hydrobiologia 167/168:275–284.

Meyer B, Martini P, Biscontin A, et al (2015) Pyrosequencing and de novo assembly of Antarctic krill (*Euphausia superba*) transcriptome to study the adaptability of krill to climate-induced environmental changes. Mol Ecol Resour 15:1460–1471. doi: 10.1111/1755-0998.12408

Miller B, von der Heyden S, Gibbons M (2012) Significant population genetic structuring of the holoplanktic scyphozoan *Pelagia noctiluca* in the Atlantic Ocean. African J Mar Sci 34:425–430. doi: 10.2989/1814232X.2012.726646

Milligan PJ, Stahl EA, Schizas N V., Turner JT (2011) Phylogeography of the copepod *Acartia hudsonica* in estuaries of the northeastern United States. Hydrobiologia 666:155–165. doi: 10.1007/s10750-010-0097-y

Minxiao W, Song S, Chaolun L, Xin S (2011) Distinctive mitochondrial genome of Calanoid copepod *Calanus sinicus* with multiple large non-coding regions and reshuffled gene order: useful molecular markers for phylogenetic and population studies. BMC Genomics 12:73. doi: 10.1186/1471-2164-12-73

Miyamoto H, Machida RJ, Nishida S (2010) Genetic diversity and cryptic speciation of the deep sea chaetognath *Caecosagitta macrocephala* (Fowler, 1904). Deep Res Part II Top Stud Oceanogr 57:2211–2219. doi: 10.1016/j.dsr2.2010.09.023

Miyamoto H, Machida RJ, Nishida S (2012) Global phylogeography of the deep-sea pelagic chaetognath *Eukrohnia hamata*. Prog Oceanogr 104:99–109. doi: 10.1016/j.pocean.2012.06.003

Moroz LL, Kocot KM, Citarella MR, et al (2014) The Ctenophore Genome and the Evolutionary Origins of Neutral Systems. Nature 510:109–114. doi: 10.1038/nature13400

Narum SR, Buerkle CA, Davey JW, et al (2013) Genotyping-by-sequencing in ecological and conservation genomics. Mol Ecol 22:2841–2847. doi: 10.1111/mec.12350

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: Identifying adaptive variation in space and time. Mol Ecol 18:3128–3150. doi: 10.1111/j.1365-294X.2009.04272.x

Nilsson B, Jepsen PM, Rewitz K, Hansen BW (2014) Expression of hsp70 and ferritin in embryos of the copepod *Acartia tonsa* (Dana) during transition between subitaneous and quiescent state. J Plankton Res 36:513–522. doi: 10.1093/plankt/fbt099

Norton EL, Goetze E (2013) Equatorial dispersal barriers and limited population connectivity among oceans in a planktonic copepod. Limnol Oceanogr 58:1581–1596. doi: 10.4319/lo.2013.58.5.1581

O'Grady JF, Hoelters LS, Swain MT, Wilcockson DC (2016) Identification and temporal expression of putative circadian clock transcripts in the amphipod crustacean *Talitrus saltator*. PeerJ 4:e2555. doi: 10.7717/peerj.2555

Ogoh K, Ohmiya Y (2004) Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracoda) with duplicate control regions. Gene 327:131–139. doi: 10.1016/j.gene.2003.11.011

Omori M, Hamner WM (1982) Patchy distribution of zooplankton: Behavior, population assessment and sampling problems. Mar Biol 72:193–200. doi: 10.1007/BF00396920

Papadopoulos LN, Peijnenburg KTCA, Luttikhuizen PC (2005) Phylogeography of the calanoid copepods *Calanus helgolandicus* and *C. euxinus* suggests Pleistocene divergences between Atlantic, Mediterranean, and Black Sea populations. Mar Biol 147:1353–1365. doi: 10.1007/s00227-005-0038-x

Papetti C, Zane L, Bortolotto E, et al (2005) Genetic differentiation and local temporal stability of population structure in the euphausiid *Meganyctiphanes norvegica*. Mar Ecol Prog Ser 289:225–235. doi: 10.3354/meps289225

Papillon D, Perez Y, Caubit X, Le Parco Y (2004) Identification of Chaetognaths as protostomes is supported by the analysis of their mitochondrial genome. Mol Biol Evol 21:2122–2129. doi: 10.1093/molbev/msh229

Papot C, Cascella K, Toullec JY, Jollivet D (2016) Divergent ecological histories of two sister Antarctic krill species led to contrasted patterns of genetic diversity in their heat-shock protein (hsp70) arsenal. Ecol Evol 6:1555–1575. doi: 10.1002/ece3.1989

Patarnello T, Papetti C, Zane L (2010) Genetics of Northern Krill (*Megantyctiphanes norvegica* Sars). In: Tarling G (ed) Biology of Northern Krill. p Vol 57

Pearman JK, Irigoien X (2015) Assessment of zooplankton community composition along a depth profile in the central Red Sea. PLoS One 10:1–14. doi: 10.1371/journal.pone.0133487

Peijnenburg KTCA, Breeuwer JAJ, Pierrot-Bults AC, Menken SBJ (2004) Phylogeography of the planktonic chaetognath *Sagitta setosa* reveals isolation in European seas. Evolution 58:1472–1487. doi: 10.1554/03-638

Peijnenburg KTCA, Fauvelot C, Breeuwer JAJ, Menken SBJ (2006) Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol Ecol 15:3319–3338. doi: 10.1111/j.1365-294X.2006.03002.x

Peijnenburg KTCA, Goetze E (2013) High evolutionary potential of marine zooplankton. Ecol Evol 3:2765–2783. doi: 10.1002/ece3.644

Peijnenburg KTCA, Van Haastrecht EK, Fauvelot C (2005) Present-day genetic composition suggests contrasting demographic histories of two dominant chaetognaths of the North-East Atlantic, Sagitta elegans and S. setosa. Mar Biol 147:1279–1289. doi: 10.1007/s00227-005-0041-2

Pereira RJ, Barreto FS, Pierce NT, et al (2016) Transcriptome-wide patterns of divergence during allopatric evolution. Mol Ecol 25:1478–1493. doi: 10.1111/mec.13579

Pereira RJ, Sasaki MC, Burton RS (2017) Adaptation to a latitudinal thermal gradient within a widespread copepod species: the contributions of genetic divergence and phenotypic plasticity. Proc R Soc B Biol Sci 284:20170236. doi: 10.1098/rspb.2017.0236

Peterson BK, Weber JN, Kay EH, et al (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One. doi: 10.1371/journal.pone.0037135

Petkeviciute E, Kania PW, Skovgaard A (2015) Genetic responses of the marine copepod *Acartia tonsa* (Dana) to heat shock and epibiont infestation. Aquac Reports 2:10–16. doi: 10.1016/j.aqrep.2015.04.001

Pett W, Ryan JF, Pang K, et al (2011) Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: Insight from mtDNA and the nuclear genome. Mitochondrial DNA 22:130–42. doi: 10.3109/19401736.2011.624611

Planque B, Graeme Hay IC, Ibanez F, Gamble JC (1997) Large scale spatial variations in the seasonal abundance of *Calanus finmarchicus*. Deep Res I 44:315–326.

Pogson GH (2016) Studying the genetic basis of speciation in high gene flow marine invertebrates. Curr Zool 62:643–653. doi: 10.1093/cz/zow093

Prokopowich CD, Gregory TR, Crease TJ (2003) The correlation between rDNA copy number and genome size in eukaryotes. Genome 46:48–50. doi: 10.1139/g02-103

Provan J, Beatty GE, Keating SL, et al (2009) High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus*. Proc Biol Sci 276:301–307. doi: 10.1098/rspb.2008.1062

Questel JM, Blanco-Bercial L, Hopcroft RR, Bucklin A (2016) Phylogeography and connectivity of the *Pseudocalanus* (Copepoda: Calanoida) species complex in the eastern North Pacific and the Pacific Arctic Region. J Plankton Res 0:1–14. doi: 10.1093/plankt/fbw025

Rahlff J, Peters J, Moyano M, et al (2017) Short-term molecular and physiological responses to heat stress in neritic copepods *Acartia tonsa* and *Eurytemora affinis*. Comp Biochem Physiol -Part A Mol Integr Physiol 203:348–358. doi: 10.1016/j.cbpa.2016.11.001

Raisuddin S, Kwok KWH, Leung KMY, et al (2007) The copepod Tigriopus: A promising marine model organism for ecotoxicology and environmental genomics. Aquat Toxicol 83:161–173. doi: 10.1016/j.aquatox.2007.04.005

Ramos AA, Weydmann A, Cox CJ, et al (2015) A transcriptome resource for the copepod *Calanus glacialis* across a range of culture temperatures. Mar Genomics 23:27–29. doi: 10.1016/j.margen.2015.03.014

Rasch EM, Lee CE, Wyngaard GA (2004) DNA-Feulgen cytophotometric determination of genome size for the freshwater-invading copepod *Eurytemora affinis*. Genome 47:559–564. doi: 10.1139/G04-014

Rawson PD, Brazeau D a, Burton RS (2000) Isolation and characterization of cytochrome c from the marine copepod *Tigriopus californicus*. Gene 248:15–22.

Reitzel AM, Herrera S, Layden MJ, et al (2013) Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and

population genomics. Mol Ecol 22:2953–2970. doi: 10.1111/mec.12228

Renaut S, Dion-Coˆté AM (2016) History repeats itself: Genomic divergence in copepods. Mol Ecol 25:1417–1419. doi: 10.1111/mec.13577

Riginos C, Crandall ED, Liggins L, et al (2016) Navigating the currents of seascape genomics: How spatial analyses can augment population genomic studies. Curr Zool 62:581–601. doi: 10.1093/cz/zow067

Romero IG, Ruvinsky I, Gilad Y (2014) Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet 13:505–516. doi: 10.1038/nrg3229.Comparative

Roncalli V, Cieslak MC, Lenz PH (2016) Transcriptomic responses of the calanoid copepod *Calanus finmarchicus* to the saxitoxin producing dinoflagellate *Alexandrium fundyense*. Nat Publ Gr 1–13. doi: 10.1038/srep25708

Ryan JF, Pang K, Schnitzler CE, et al (2013) The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. Science (80-) 342:1242592. doi: 10.1126/science.1242592

Saborowski R, Buchholz F (2002) Metabolic properties of Northern krill, *Meganyctiphanes norvegica*, from different climatic zones. I. Respiration and excretion. Mar Biol 140:547–556. doi: 10.1007/s00227-001-0730-4

Sales G, Deagle BE, Calura E, et al (2017) KrillDB: A de novo transcriptome database for the Antarctic krill (*Euphausia superba*). PLoS One 12:1–12. doi: 10.1371/journal.pone.0171908

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. Nat Rev Genet 14:807–20. doi: 10.1038/nrg3522

Schindel DE, Miller SE (2005) DNA barcoding a useful tool for taxonomists. Naure 435:17–17.

Schlötterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. Nat Rev Genet 15:749–763. doi: 10.1038/nrg3803

Schoville SD, Barreto FS, Moy GW, et al (2012) Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. BMC Evol Biol 12:170. doi: 10.1186/1471-2148-12-170

Schunter C, Carreras-Carbonell J, MacPherson E, et al (2011) Matching genetics with oceanography: Directional gene flow in a Mediterranean fish species. Mol Ecol 20:5167–5181. doi: 10.1111/j.1365-294X.2011.05355.x

Shao Z, Graf S, Chaga OY, Lavrov D V. (2006) Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): A linear DNA molecule encoding a putative DNA-dependent DNA polymerase. Gene 381:92–101. doi: 10.1016/j.gene.2006.06.021

Shen X, Wang H, Ren J, et al (2010) The mitochondrial genome of *Euphausia superba* (Prydz Bay) (Crustacea: Malacostraca: Euphausiacea) reveals a novel gene arrangement and potential molecular markers. Mol Biol Rep 37:771–784. doi: 10.1007/s11033-009-9602-7

Shen X, Wang H, Wang M, Liu B (2011) The complete mitochondrial genome sequence of *Euphausia pacifica* (Malacostraca: Euphausiacea) reveals a novel gene order and unusual tandem repeats. Genome 54:911–22. doi: 10.1139/g11-053

Sherman CDH, Lotterhos KE, Richardson MF, et al (2016) What are we missing about marine invasions? Filling in the gaps with evolutionary genomics. Mar Biol 163:1–24. doi: 10.1007/s00227-016-2961-4

Siegel V, Watkins JL (2016) Distribution, Biomass and Demography of Antarctic Krill, *Euphausia superba*. In: Volker S (ed) Biology and Ecology of Antarctic Krill. Springer International Publishing, pp 21–100

Skjoldal HR, Wiebe PH, Postel L, et al (2013) Intercomparison of zooplankton (net) sampling systems: Results from the ICES/GLOBEC sea-going workshop. Prog Oceanogr 108:1–42. doi: 10.1016/j.pocean.2012.10.006

Smolina I (2015) *Calanus* in the North Atlantic: species identification, stress response, and population genetic structure. PhD Aquat Biosci 15:64.

Smolina I, Harmer R, Lindeque P, Hoarau G (2016) Reduced up-regulation of gene expression in response to elevated temperatures in the mid-Atlantic population of *Calanus finmarchicus*. J Exp Mar Bio Ecol 485:88–93. doi: 10.1016/j.jembe.2016.09.003

Smolina I, Kollias S, Møller E, et al (2015) Contrasting transcriptome response to thermal stress in two key zooplankton species, *Calanus finmarchicus* and *C. glacialis*. Mar Ecol Prog Ser 534:79–93. doi: 10.3354/meps11398

Smolina I, Kollias S, Poortvliet M, et al (2014) Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. Mol Ecol Resour 14:1072–1079. doi: 10.1111/1755-0998.12241

Stapley J, Reger J, Feulner PGD, et al (2010) Adaptation genomics: The next generation. Trends Ecol Evol 25:705–712. doi: 10.1016/j.tree.2010.09.002

Stopar K, Ramšak A, Trontelj P, Malej A (2010) Lack of genetic structure in the jellyfish *Pelagia noctiluca* (Cnidaria: Scyphozoa: Semaeostomeae) across European seas. Mol Phylogenet Evol 57:417–428. doi: 10.1016/j.ympev.2010.07.004

Sun C, Wyngaard G, Walton DB, et al (2014) Billions of basepairs of recently expanded, repetitive sequences are eliminated from the somatic genome during copepod development. BMC Genomics 15:186. doi: 10.1186/1471-2164-15-186

Sun C, Zhao Y, Li H, et al (2015) Unreliable quantitation of species abundance based on high-throughput sequencing data of zooplankton communities. Aquat Biol 24:9–15. doi: 10.3354/ab00629

Tarrant AM, Baumgartner MF, Hansen BH, et al (2014) Transcriptional profiling of reproductive development, lipid storage and molting throughout the last juvenile stage of the marine copepod *Calanus finmarchicus*. Front Zool 11:91. doi: 10.1186/s12983-014-0091-8

Thabet AA, Maas AE, Saber SA, Tarrant AM (2017) Assembly of a reference transcriptome for the gymnosome pteropod *Clione limacina* and profiling responses to short-term CO_2 exposure. Mar Genomics 34:39–45. doi: 10.1016/j.margen.2017.03.003

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21:3907–3930. doi: 10.1111/j.1365-294X.2012.05664.x

Toullec JY, Corre E, Bernay B, et al (2013) Transcriptome and Peptidome Characterisation of the Main Neuropeptides and Peptidic Hormones of a Euphausiid: The Ice Krill, *Euphausia crystallorophias*. PLoS One. doi: 10.1371/journal.pone.0071609

Tsagkogeorga G, Cahais V, Galtier N (2012) The population genomics of a fast evolver: High levels of diversity, functional constraint, and molecular adaptation in the tunicate *Ciona intestinalis*. Genome Biol Evol 4:740–749. doi: 10.1093/gbe/evs054

Unal E, Bucklin A (2010) Basin-scale population genetic structure of the planktonic copepod *Calanus*

finmarchicus in the North Atlantic Ocean. Prog Oceanogr 87:175–185. doi: 10.1016/j.pocean.2010.09.017

Voolstra CR, W??rheide G, Lopez J V. (2017) Advancing genomics through the Global Invertebrate Genomics Alliance (GIGA). Invertebr Syst 31:1–7. doi: 10.1071/IS16059

Wang K, Omotezako T, Kishi K, et al (2015) Maternal and zygotic transcriptomes in the appendicularian, *Oikopleura dioica*: novel protein-encoding genes, intra-species sequence variations, and trans-spliced RNA leader. Dev Genes Evol 225:149–159. doi: 10.1007/s00427-015-0502-7

Wang S, Meyer E, McKay JK, Matz M V (2012) 2b-RAD: a simple and flexible method for genome-wide genotyping. Nat Methods 9:808–810. doi: 10.1038/nmeth.2023

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. J Hered 89:438–450. doi: 10.1093/jhered/89.5.438

Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: How can we do it better? Fish Fish 9:423–449. doi: 10.1111/j.1467-2979.2008.00303.x

Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. Mar Ecol Prog Ser 393:1–12. doi: 10.3354/meps08287

Wei S, Li P, Yang M, et al (2016) The mitochondrial genome of the pelagic chaetognath, *Pterosagitta draco*. Mitochondrial DNA Part B 1:515–516. doi: 10.1080/23802359.2016.1197055

Whitehead A (2012) Comparative genomics in ecological physiology: toward a more nuanced understanding of acclimation and adaptation. J Exp Biol 215:884–891. doi: 10.1242/jeb.058735

Wiebe PH, Harris RP, St. John MA, et al (2009) BASIN: Basin-scale Analysis, Synthesis, and INtegration. Science Plan and Implementation Strategy. GLOBEC Report 27: iii, 43pp.

Wiebe PH, Bucklin A, Madin L, et al (2010) Deep-sea sampling on CMarZ cruises in the Atlantic Ocean - an Introduction. Deep Res Part II Top Stud Oceanogr 57:2157–2166. doi: 10.1016/j.dsr2.2010.09.018

Wiebe PH, Lawson GL, Lavery AC, et al (2013) Improved agreement of net and acoustical methods for surveying euphausiids by mitigating avoidance using a net-based LED strobe light system. ICES J Mar Sci 70:650–664. doi: 10.1093/icesjms/fsr005

Wiebe PH (1968) Plankton patchiness: effects on repeated net tows. Limnol Oceanogr 13:315–321. doi: 10.4319/lo.1968.13.2.0315

Wyngaard GA, McLaren IA, White MM, Sévigny JM (1995) Unusually high numbers of ribosomal RNA genes in copepods (Arthropoda: Crustacea) and their relationship to genome size. Genome 38:97–104. doi: 10.1139/g95-012

Wyngaard GA, Rasch EM (2000) Patterns of genome size in the copepoda. Hydrobiologia 417:43–56. doi: 10.1023/A:1003855322358

Wyngaard GA, Rasch EM, Connelly BA (2011) Unusual augmentation of germline genome size in *Cyclops kolensis* (Crustacea, Copepoda): Further evidence in support of a revised model of chromatin diminution. Chromosom Res 19:911–923. doi: 10.1007/s10577-011-9234-3

Yang EJ, Ha HK, Kang S-H (2014) Microzooplankton community structure and grazing impact on major phytoplankton in the Chukchi sea and the western Canada basin, Arctic ocean. Deep Sea Res Part II Top Stud Oceanogr 120:91–102. doi: 10.1016/j.dsr2.2014.05.020

Zane L, Ostellari L, Maccatrozzo L, et al (1998) Molecular Evidence for Genetic Subdivision of Antarctic

Krill (*Euphausia superba* Dana) Populations. Proc Biol Sci 265:2387–2391.

Zane L, Ostellari L, Maccatrozzo L, et al (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica*: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar Biol 136:191–199. doi: 10.1007/s002270050676

Zane L, Patarnello T (2000) Krill: a possible model for investigating the effects of ocean currents on the genetic structure of a pelagic invertebrate. Can J Fish Aquat Sci 57:16–23. doi: 10.1139/f00-166

Figures

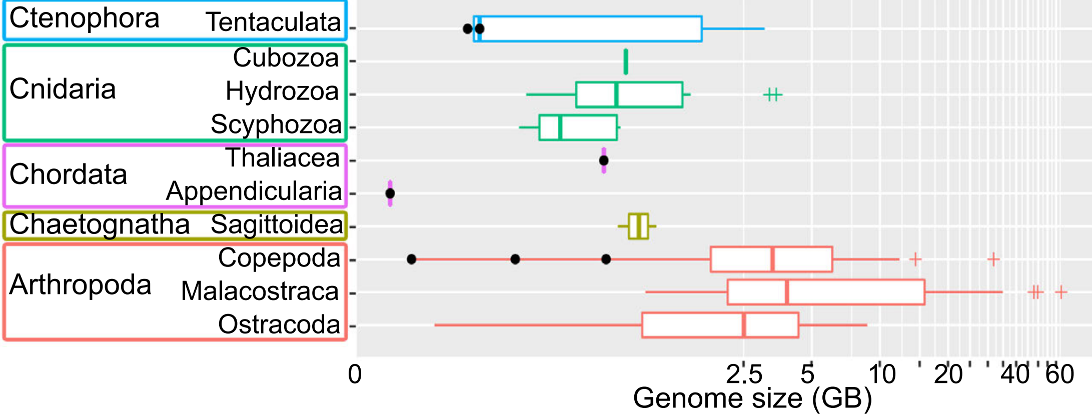


Fig.1

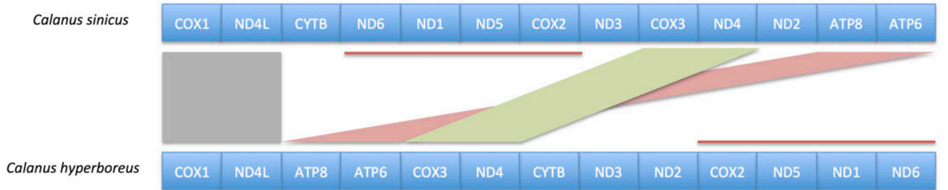


Fig. 2

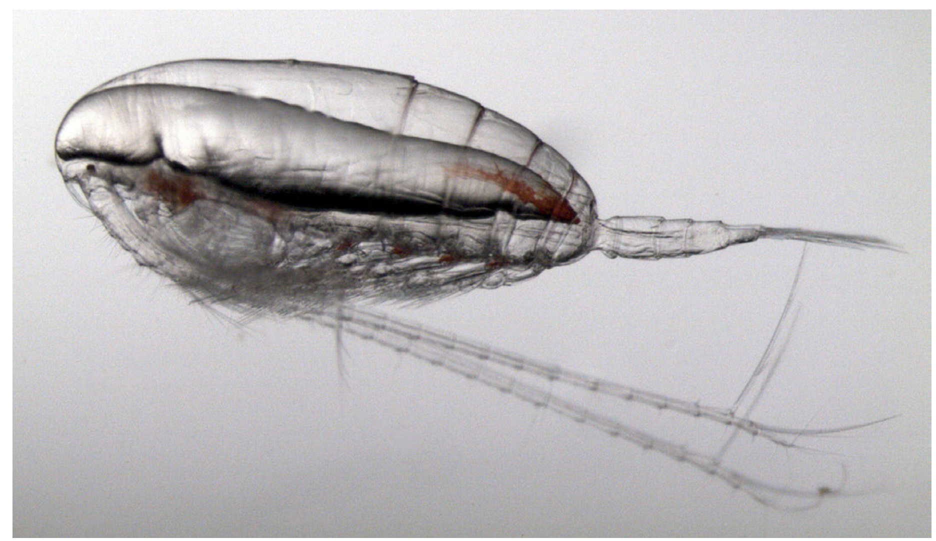


Fig. 3

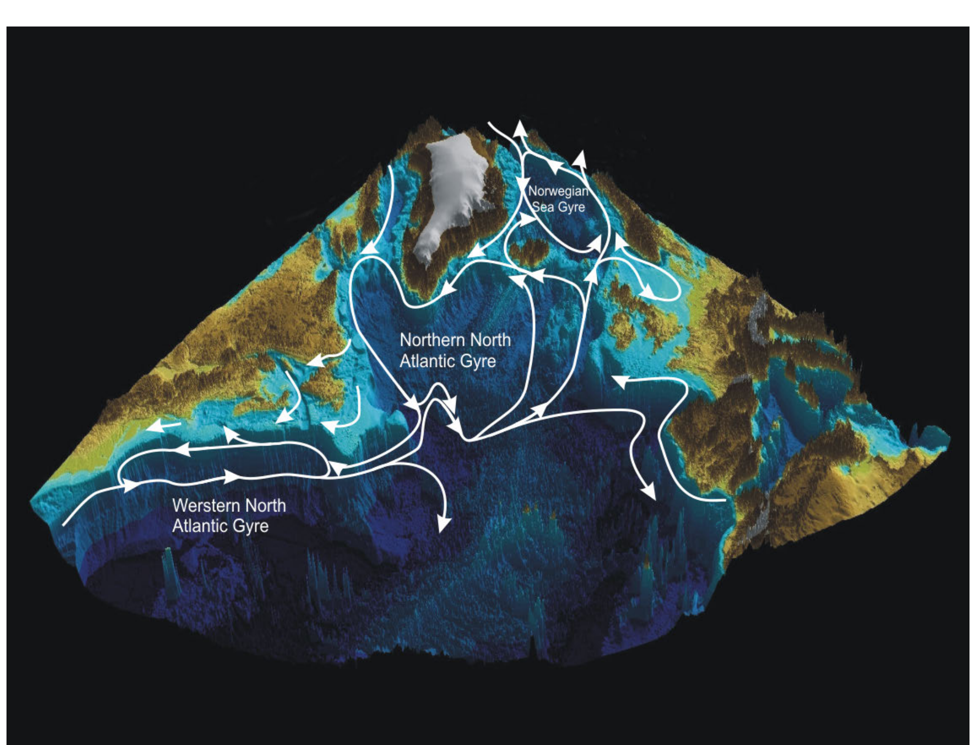


Fig. 4

Fig. 5

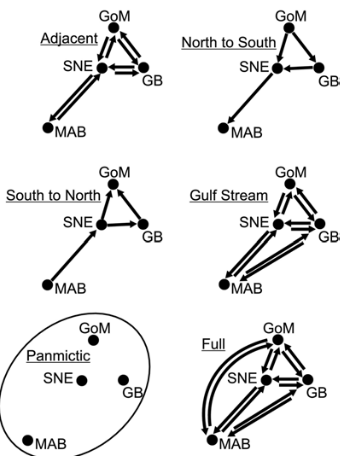


Fig. 6

Fig. 7

Fig. 8

Fig. 9

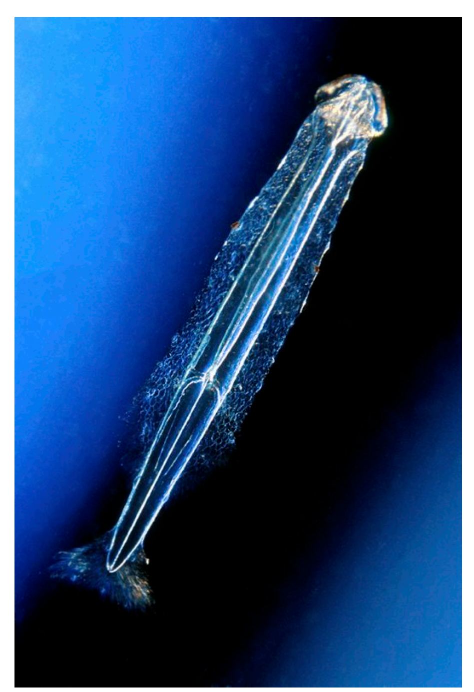


Fig. 10

Fig. 11

Tables

Table 1.

	Tunicates			Ctenophores		Copepods			
	Oikopleura dioica	*Oikopleura dioica*	*Salpa thompsoni*	*Pleurobrachia bachei*	*Mnemiopsis leidyi*	*Eurytemora affinis*	*Oithona nana*	*Tigriopus kingsejongensi*	*Tigriopus californicus*
	Main assem	Allelic assem							
Assembly Name	ASM20953v1	ASM20955v1	Salp Genome 1.0	P.bachei_draft_genome_v.1.1	GCA_000226015.1	Eaff_1.0	O. nana v.1.0	NA	TCALIF_v1.0
Estimated Genome Size (MB)	68,46	allelic assembly	602[a]	170	150	616.14[b]	85	298	245[d]
Assembly Size	70 471 451	45 141 193	318 747 957	156 121 975	155 865 547	494 890 867	85 010 107	305 712 242	184 634 130
Predicted protein coding Genes	18 020	18 020	13 186	19 523	16 548	29 783	15 359	12 772	14 536
Coverage	14X	N/A[c]	20X	200X	160X	75X	N/A[c]	65X	N/A[c]
Number of scaffolds	1 260	4 196	478 281	21 979	5 100	6 899	4 626	27 823	2 365
Length of N50 scaffold (bp)	395 387	21 890	934	20 628	187 314	862 645	400 614	159 218	298 012
Number of N50 scaffold (L50)	35	478	79 492	1 646	242	163	60	N/A[c]	180
Number of contigs	5 917	6 678	590 021	38 864	24 884	122 625	7 437	48 368	26 175
Length of N50 contig (bp)	24 932	10 847	636	6 132	11 936	5 738	38 620	17 566	14 799
Number of N50 contig	718	985	136 534	6 078	3 653	19 338	463	N/A[c]	3 352
Length of gaps (bp)	3 938 358	2 655 217	14 945 692	19 276 734	5 525 119	107 316 113	2 943 785	10 474 460	N/A[c]
References	Denoeud et al. (2010)	Denoeud et al. (2010)	Jue et al. (2016)	Moroz et al. (2014)	Ryan et al. (2013)	Eyun et al. (2017)	Madoui et al. (2017)	Kang et al. (2017)	https://i5k.nal.usda.gov/Tigriopus_californicus

[a] genome size estimate independent of assembly (Jue et al. (2016))
[b] genome size estimate independent of assembly (Rasch et al. (2004))
[c] not available/not provided
[d] genome size estimate independent of assembly (Wyngaard and Rasch 2000)

Table 1

Table 2.

Taxon and Species	Citation	Length (bp)
Copepoda		
Calanus hyperboreus	Kim et al., 2013	17 910
Calanus sinicus	Minxiao et al., 2011	>20,460
Paracyclopina nana	Ki et al., 2009	15 981
Tigriopus californicus	Burton et al., 2007	14 600
Tigriopus japonicus	Machida et al., 2002	14 628
Tigriopus sp.	Jung et al., 2006	14 301
Euphausiacea		
Euphausia pacifica	Shen et al., 2011	16 898
Euphausia superba	Shen et al., 2010	>15,498
Ostracoda		
Vargula hilgendorfii	Ogoh & Ohmiya, 2004	15 923
Amphipoda		
Onisimus nanseni	Ki et al., 2010	14 734
Decapoda		
Acetes chinensis	Kim et al., 2012	15 740
Cnidaria		
Aurelia aurita	Shao et al., 2006	16 937
Cassiopea frondosa	Kayal et al., 2011	15 949
Chrysaora quinquecirrha	Hwang et al., 2014	16 775
Ctenophora		
Mnemiopsis leidyi	Pett et al., 2011	10 000
Pleurobrachia bachei	Kohn et al., 2012	11 016
Chaetognatha		
Sagitta decipiens	Miyamoto et al., 2010	11 121
Sagitta enflata	Miyamoto et al., 2010	12 631
Sagitta ferox	Li et al., 2016	12 153
Sagitta nagae	Miyamoto et al., 2010	11 459
Paraspadella gotoi	Helfenbein et al., 2004	11 423
Pterosagitta draco	Wei et al., 2016	10 426
Spadella cephaloptera	Papillon et al., 2004	11 905

Table 2

Table 3.

Phylum and Species	BioProject	Contig Total	Contig Max Length	Contigs Total Length	Contigs Annotated	Transcripts	N50	Genes	Citation
Cnidaria									
Alatina alata	PRJNA312373	31 737	32 591	48 508 802	No	31 776	2 545	20 173	Ames et al. (2016)
Rhopilema esculentum	PRJNA318143	148 857	30 742	121 470 903	No	NA	NA	NA	Chongbo and Yunfeng (Dir Sub)
Aurelia aurita	PRJNA252562	252 170	46 960	180 188 094	No	24 264	1 761	10 285	Brekham et al. (2015)
Ctenophora									
Mnemiopsis leidyi	PRJNA344880	140 842	29 348	137 638 938	No	NA	NA	NA	Sanchez Alvarado, Gotting and Ross (Dir Sub)
Arthropoda: Copepoda									
Acartia fossae		100 383	8 174		No		769		Eyun et al. (2017)
Calanus finmarchicus	PRJNA236983	28 954	2 945	10 223 122	No	251 042	354	13 057	Smolina et al. (2014)
Calanus finmarchicus	PRJNA236528	206 012	23 068	205 455 659	Yes		1 418		Lenz et al. (2014)
Calanus finmarchicus	PRJNA231164	241 140	25 048	160 760 719	No				Tarrant et al. (2014)
Calanus glacialis	PRJNA237014	36 880	4 021	15 748 490	No	242 602	471	18 387	Smolina et al. (2014)
Calanus glacialis	PRJNA274584	54 344	7 507	33 214 362	No	16 998	620	16 998	Ramos et al. (2014)
Calanus sinicus		69 751				69 751	1 127	43 417	Yang et al. (2014)
Calanus sinicus			3 923		No	29 458	513		Eyun et al. (2017)
Eucyclops serrulatus	PRJNA231234	51 528	16 342	36 645 141	No				Cattonaro (Dir Sub)
Eurytemora affinis	PRJNA278152	107 445	26 685	142 143 154	No	29 783			Monroe (Dir Sub)
Eurytemora affinis	PRJNA242763	138 088	23 627	143 733 589	Yes				Almada and Tarrant (Dir Sub)
Eurytemora affinis		88 104	26 685						Eyun et al. (2017)
Paracyclopina nana	PRJNA268783	60 687	27 858	95 849 484	Yes	67 179	4 178	12 474	Lee et al. (2015)
Pseudocalanus acuspes	PRJNA296544	207 302	12 713	59 236 626	Yes	69 555	1 348	28 879	De Wit et al. (2016)
Tigriopus kingsejongensis	PRJNA283925	38 250	7 809	36 497 199	Yes				Lee (Dir Sub)
Tigriopus kingsejongensis			23 942	28 850 726		40 172	1 093	12 772	Kang et al. (2017)
Tigriopus californicus	PRJNA263967	12 067	13 452	14 966 851	No				Baretto et al. (2010)
Tigriopus californicus	PRJNA263967	12 075	13 452	14 902 878	No				Baretto et al. (2010)
Tigriopus californicus		106 317	27 644	NA	Yes	106 317	2 837	12 573	Periera et al. (2016)
Tigriopus californicus		60 840	8 614				1 510		Eyun et al. (2017)
Tigriopus japonicus	PRJNA274317	54 758	23 769	82 981 758	Yes		3 565		Kim et al. (2015)
Arthropoda: Euphausiacea									
Euphausia superba			11 127		Yes	15 347	520	7 942	Meyer et al. (2015)
Euphausia superba		22 177	8 515		Yes			5 563	Clark et al. (2011)
Euphausia superba		133 962		129 183 922	Yes		1 294	27 928	Sales et al 2017
Euphasia crystallorophias		42 632					8 341		Toullec et al. (2013)
Meganyctiphanes norvegica	PRJNA324094	405 497	26 644	222 530 071	No	NA	NA	NA	Maas and Blanco Bercial (Dir Sub)
Arthropoda: Amphipoda									
Talitrus saltator	PRJNA297565	156 706	22 032	151 674 147	Yes		968		O'Grady et al. (2016)
Arthropoda: Mysidacea									
Neomysis awatschensis	PRJNA287057	22 141	10 398	14 999 154	Yes	22 141	801		Kim et al. (2016)
Mollusca: Pteropoda									
Clio pyramidata	PRJNA231010	45 735				45 735	852	30 800	Maas et al. (2015)
Clione limacina	PRJNA314884	477 401	30 190	258 267 445	Yes	300 994	816	181 879	Thabet et al. (2017)
Limacina antarctica	PRJNA295792	81 226	7 935	59791880	No	402 273	500	81 229	Johnson and Hoffman (2016)
Limacina helicina	PRJNA386290	53 121	12 358	31 790 000	Yes		796		Koh et al. 2015
Urochordata: Tunicata									
Oikopleura dioica	PRJNA269316	54 949	23 096	66 526 340	No				Wang et al. (2015)
Oikopleura dioica	PRJNA269317	86 898		70 800 000		57 962	1 806	16 423	Wang et al. (2015)
Salpa thompsoni	PRJNA279245	217 849	30 785	151 741 986	No	216 931	1 163	26 413	Jue et al. (2016); Batta Lona et al. (2017)

Table 3