PhD in Aquatic Biosciences no. 15 (2015)

UNIVERSITY OF NORDLAND

Irina Smolina

Calanus in the North Atlantic: species identification, stress response, and population genetic structure

Calanus in the North Atlantic: species identification, stress response, and population genetic structure

Irina Smolina

A thesis for the degree of
Philosophiae Doctor (PhD)

PhD in Aquatic Biosciences (2015)
Faculty of Biosciences and Aquaculture

PhD in Aquatic Biosciences no. 15 (2015)

Irina Smolina
Calanus in the North Atlantic: species identification, stress response, and population genetic structure

ISBN: 978-82-93165-14-9

Print: Trykkeriet UiN

University of Nordland
N-8049 Bodø
Tel: +47 75 51 72 00
www.uin.no

Preface

This dissertation is submitted in partial fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Faculty of Biosciences and Aquaculture (FBA), University of Nordland (UiN), Bodø, Norway. The different studies compiled in this dissertation represent original research carried out over a period of three years, from 01.08.2012 to 31.07.2015, as part of the project "*Calanus* in the North Atlantic: species distribution and genetic population structure in space and time" funded by the Norwegian Research Council (HAVKYST 216578, 234356) to Professor Galice Hoarau and part of FP7 EURO-BASIN (Grant Agreement: 264 933).

The project team included the following members:

Irina Smolina, MSc, UiN, PhD student

Galice Hoarau, Professor, UiN, primary supervisor

Ann Bucklin, Professor, Department of Marine Sciences, University of Connecticut, co-supervisor

Irina Smolina

Bodø, July 2015

Acknowledgements

"Curiouser and curiouser!" Cried Alice (she was so much surprised, that for the moment she quite forgot how to speak good English)."

— Lewis Carroll, Alice's Adventures in Wonderland & Through the Looking-Glass

I should admit that many times during this *Calanus* project I felt pretty much as Carroll´s Alice, wondering about a new unexpected turn or mystery. Enjoying the final result, I would like to say many thanks to all people who were involved in the project, helped to fulfil it the right way, and supported me throughout.

First, I would like to gratefully thank my primary supervisor, Professor Galice Hoarau, for continuation with enthralling challenges since the start of my master research project; for inspiration, inexhaustible ideas and interesting discussions; supervising me and the project and mentoring "How it works in research world". It was quite fun and exciting to see an evolution of this project from "simple and boring, with 1000s of PCR plates" into the project of cutting-edge NGS approaches in marine ecology. And particular thanks for reminders to keep a balanced life and for introduction to scuba diving and powder skiing! Great moments to remember!

I am sincerely grateful to my co-supervisor, Ann Bucklin, for firing my enthusiasm about the mysterious *Calanus* genome, stimulating conversation, encouragement and great help at the last stage of the dissertation; all of these, despite being on the other side of North Atlantic Ocean. I´m particularly grateful to Ann and Peter Wiebe for being my allies for several long weeks during the Trans-Atlantic cruise, sharing experience and knowledge, space and edible food supplies.

Many thanks to my senior colleagues, who, knowingly or unknowingly, contributed to the success of this research: Slawek Kwasniewski, for giving a good taxonomic base in the world of so similar copepod species; Petra Lenz for hosting me heartily in her group in University of Hawaii at Manoa (US) and I´m grateful to her and Vittoria Roncalli for productive discussions on RNA-seq and *Calanus* physiology, and for a very stimulating and creative atmosphere; Torkel G Nielsen, Pennie Lindeque and Eva F Møller for assistance with sampling and stress experiments, and for sharing ecological

point of view; Jim Coyer and Leo Blanco-Bercial for always being happy to help and encourage. And many thanks to researchers and crews of the research vessels, who participated in the cruises RV “Johan Hjort” 201208, RV “Helmer Hanssen” study course cruise 201209 and RV “GO Sars” EUROBASIN 201305, and helped in various ways.

I gratefully thank to friends and all colleagues at FBA, from previous and current times, for advises, technical and administrative assistance. Particularly, thanks to Maren, Cindy, Spyros & Marol, Jorge, Arvind & Priya, Carlo, Binoy, Kazue, Teshome, Marloes, Tor Erik, Renate, Sünnje, Katrin & Hennig, Hin, Andrea, Micha, Lokesh, Anya, Ira, Marina, Boris, Marvin, Martina, Melisa, Mark and Jeanett for making this time lively, fun, and enjoyable at and out of work. I would like to single out and say special thanks to: Spyros, a NGS guru, who taught me a lot; Lokesh, for sharing the burden of all-alone Ion Torrent sequencing; Marvin, inheritor of the *Calanus* project, who brought fresh enthusiasm when mine was out of order; and Annushka, my musicroommate and comrade in happiness, frustration, and unexplained.

Many thanks to my friends in Tomsk and elsewhere, for keeping in touch and empathy, due to whom I always regret that technological progress still does not employ teleportation.

My heartily thanks to my family for great support throughout time, for warm welcome in Bodø, for Italian sun during winter darkness, for charming Weihnachten´s, and for many more. Particularly, I am sincerely grateful to my husband Alexander for being with me in happy and gloomy moments, supporting, motivating and loving; and to my parents for encouraging and standing firmly for me.

Table of contents

List of figures and tables

Abbreviations

AFLP	Amplified fragment length polymorphism
Ampli-seq	Amplicon sequencing
BLAST	Basic local alignment search tool
Cap-seq	Capture sequencing
cDNA	Complementary DNA
COI	Cytochrome oxidase subunit 1
Cyt b	Cytochrome b
ddRAD	Double digest restriction-site associated DNA
DNA	Deoxyribonucleic acid
EST	Expressed sequence tag
GBS	Genotyping-by-sequencing
GO	Gene ontology
GWAS	Genome-wide association study
HSP	Heat shock protein
InDel	Insertion-deletion polymorphism
ITS	Internal transcribed spacer
KEGG	Kyoto encyclopedia of genes and genomes
mtDNA	Mitochondrial DNA
NGS	Next generation sequencing
nucDNA	Nuclear DNA
PCR	Polymerase chain reaction
QTL	Quantitative trait locus
RAD-seq	Restriction-site associated DNA sequencing
RAPD	Random amplified polymorphism detection
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RRL	Reduced representation library
rRNA	ribosomal RNA
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
qPCR	Quantitative polymerase chain reaction

Glossary of next generation sequencing terms

Adapter	Short double stranded DNA molecule, which is used to link DNA fragment of interest with a sequencing primer for next generation sequencing
Annotation	Identification of functional role of genome or transcriptome sequence
Barcode (index)	Short sequence identifier (about 6 nucleotides) for individual labelling (barcoding) of sequencing libraries
Cap-seq (capture sequencing)	Technique for next generation sequencing of selected genomic regions via hybridization of fragmented DNA with designed capture baits
Contig	Consensus sequence of overlapping reads from high throughput sequencing. Depending on the source material for sequencing contig could represent genome locus or transcript
Coverage (depth)	Number of reads per locus
ddRAD-seq (double digest restriction-site associated DNA sequencing)	Technique for next generation sequencing of genome fraction via reduction of DNA by digestion with two restriction enzymes and tight size selection of resultant fragments that are flanked by both enzymes
***De novo* assembly**	Aligning and merging of NGS reads into longer original DNA sequence without previous knowledge of the sequence. Assembled sequences are contigs, while software tools performing this task are assemblers
Gene ontology (GO)	Structured and controlled vocabularies and classifications of gene function across species representing gene product properties on three different levels (molecular function, cellular component and biological process)
Genome-wide association study (GWAS)	Trait mapping studies that rely on a statistical test to determine associations between sequence variants and a given phenotype in natural populations
Genotyping-by-sequencing (GBS)	Set of approaches that allows polymorphism discovery and genotyping via sequencing of a genome fraction using various reduced-representation protocols
Library	Collection of fragmented DNA or cDNA belonging to one or more individuals that is modified to serve as a template for subsequent next generation sequencing
Mapping	Alignment of NGS reads to a known reference sequence or *de novo* assembly
Multiplexing	Combination of barcoded libraries for next generation sequencing
Next generation sequencing (NGS)	Number of different sequencing technologies allowing for high-throughput and massively parallel sequencing

Pool-seq	Next generation sequencing of pools of individuals without barcodes but belonging to homogeneous group (population or treatment)
Quantitative trait locus (QTL) mapping	Statistical analysis designed to investigate the genetic basis of continuous phenotypic variation
RAD-seq (restriction-site associated DNA sequencing)	Technique for next generation sequencing of genome fraction via reduction of DNA by digestion with one restriction enzymes and sequencing of short regions that flank the restriction sites
Read	Short nucleotide sequence obtained from sequencing with an NGS platform. Depending on platform and protocol read could be a single-end (when library fragment was sequenced from one end) or pair-end (when library fragment was sequenced from both ends)
Reduced representation library (RRL)	Set of approaches to prepare library for NGS that contains only a fraction of genome
RNA-seq (whole transcriptome shotgun sequencing)	Next generation sequencing of cDNA in order to get information about samples' RNA content in terms of nucleotide sequences and expression levels
Transcriptome	All RNA molecules produced in cell or organism at any one time, including mRNA, rRNA, tRNA and non-coding RNA
Trimming	Removal of poor quality sequence, poly-A tails, adaptors, barcodes and sequencing primers
Variant calling	Computational identification of locus-specific sequence polymorphism

List of papers

Paper I **Smolina I**, Kollias S, Poortvliet M, Nielsen TG, Lindeque P, Castellani C, Møller EF, 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 14:1072-1079.

Paper II[*] Nielsen TG, 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.

Paper III **Smolina I**, Kollias S, Møller EF, Lindeque P, Sundaram AYM, Fernandes JMO, Hoarau G (2015)
Contrasting transcriptome response to thermal stress in two key zooplankton species, *Calanus finmarchicus* and *C. glacialis*.
Marine Ecology Progress Series.
Accepted manuscript, doi: 10.3354/meps11398.

Paper IV **Smolina I**, Kopp M, Jüterbock A, Bucklin A, Hoarau G
Pool-ddRAD-seq reveals large-scale population differentiation of the planktonic copepod *Calanus finmarchicus* in the North Atlantic.
Manuscript.

[*] Paper II is an additional contribution to the three required papers for this dissertation, due to the fact that the paper has already been used as a chapter for PhD thesis by one of the co-authors. Nevertheless, work performed by the mentioned co-author and the author of this dissertation was independent, not overlapping, and of sufficient contribution. In addition, the paper is tightly connected to the dissertation and contributes to more complete presentation of the achieved results during PhD period of the author.

Abstract

Copepods of the genus *Calanus* are predominant in the zooplankton biomass of the North Atlantic and the Arctic where they play a key role in marine ecosystems both as main primary consumers and as prey species for fish, birds and marine mammals. Over the last 50 years, the four *Calanus* species co-occurring in the North Atlantic have shown climate-driven changes in distribution, resulting in a decrease of abundance of ecologically and commercially important fishes. Despite the plethora of studies published on *Calanus*, there is still very limited knowledge at species, population and genome levels, impeding our understanding of their ability to cope with changing climate. Therefore, the main aim of this dissertation is to address crucial knowledge gaps in *Calanus* species in the North Atlantic, specifically, unbiased identification of *Calanus* species, transcriptome response to temperature component of climate change, and genetic differentiation between populations of *C. finmarchicus*.

Challenges in *Calanus* species identification were clarified with the development of a panel of 12 nuclear insertion/deletion markers based on both genome and transcriptome sequences of *C. finmarchicus* and *C. glacialis*. All markers show species-specific amplicon length, allowing easy and robust molecular identification of *C. finmarchicus*, *C. glacialis*, and their hybrids, as well as co-occurring *C. helgolandicus*, *C. hyperboreus* and *C. marshallae*. These markers were also used to validate a new morphological criterion – redness – that can easily separate live females of *C. finmarchicus* and *C. glacialis* by the red pigmentation of the antenna and somites.

Temperature-mediated changes in gene expression in response to realistic thermal stresses (at +5 °C, +10 °C and +15 °C) for 4 hours and 6 days were investigated in *C. finmarchicus* and *C. glacialis* using whole transcriptome profiling. *C. finmarchicus* showed a strong response to low and high temperatures and long duration of stress, initiating up-regulation of genes related to protein folding, transcription, translation, and metabolism. In contrast, *C. glacialis* displayed only low-magnitude changes in gene expression. Differences in the thermal responses between the two species, particularly the weak molecular stress response in *C. glacialis*, are in line with laboratory and field observations and suggest a vulnerability of *C. glacialis* to climate change in the Arctic.

Population genetic structure of *C. finmarchicus* was investigated using a novel approach for large-scale genome-wide genotyping, pooled double digest restriction-site-associated DNA sequencing (Pool-ddRAD-seq), as well as more traditional approach, microsatellite markers. Pool-ddRAD-seq outperformed microsatellite-based method and detected both between- and within basin differentiation in the North Atlantic, additionally suggesting some degree of annual retention of *C. finmarchicus* within a region.

Overall, these findings contribute to a better understanding of the ecology and genetics of *Calanus* species and their potential to cope with climate change. This knowledge can now be implemented for managing of *Calanus* as an important biological resource, as well as for modeling ecological processes at individual, community and ecosystem levels. This dissertation also illustrates how next generation sequencing technologies can be used in non-model species to investigate a wide range of questions.

1. Introduction

1.1. *Calanus* in the North Atlantic

Planktonic marine copepods of the genus *Calanus* (Crustacea: Copepoda: Calanoida: Calanidae) are likely the most abundant animals in the world (Bucklin et al. 1996). Over their distribution ranges, the 14 species that comprise *Calanus* genus contribute significantly to the biomass (up to 70 %) and largely dominate in the ocean (Conover 1988; Head et al. 2003; Kwasniewski et al. 2003; Richardson et al. 2003; Uye 2000). Furthermore, many ecologically and commercially important fishes depend on *Calanus* species as a food source, including Atlantic and polar cod, sardine, anchovy and herring (e.g., Gislason and Astthorsson 2002; Meng 2003; Søreide et al. 2008; Uye 2000), as well as planktivorous invertebrates, birds (e.g., little auk) and marine mammals (Michaud and Taggart 2007; Skjoldal et al. 2004; Varpe et al. 2005; Weslawski et al. 1999). Thus, *Calanus* species are playing a key role in the energy transfer from the primary production level to higher trophic levels, as well as in the carbon cycle in the ocean (Beaugrand 2009; Falk-Petersen et al. 2007; Nuwer et al. 2008; Pasternak et al. 2002). The central position of *Calanus* in the ecosystem is reflected in over 25,000 publications (hits for a search in Google Scholar, accessed in May 2015), with almost half devoted to *C. finmarchicus*. Moreover, *C. finmarchicus* has been the target species of several basin-scale research programs, including investigations of migration between oceanic and shelf seas off Northwest Europe (ICOS: e.g., Heath et al. 1999), Trans Atlantic studies of *Calanus finmarchicus* (TASC: e.g., Tande and Miller 2000), the Global Ocean Ecosystem Dynamics program (GLOBEC: e.g., Gifford et al. 2010), and the European Basin-scale Analysis, Synthesis and Integration program (EURO-BASIN: e.g., Melle et al. 2014a,b).

Boreal *Calanus finmarchicus* (Gunnerus, 1770) has its core distribution in the North Atlantic, where three other *Calanus* species co-occur: *C. helgolandicus* (Claus, 1863), a temperate species in the coastal and continental shelf environment; two Arctic species, *C. glacialis* Jaschnov, 1955 and *C. hyperboreus* (Krøyer, 1838), respectively dominating on-shelf and off-shelf waters (Beaugrand et al. 2002; Bonnet et al. 2005; Conover 1988; Fleminger and Hulsemann 1977; Kosobokova and Hirche 2001) (Fig. 1

and Fig. 2). The core distribution of each species appears to be shaped by their distinct temperature niches: maximum abundances of *C. hyperboreus* and *C. glacialis* are at temperatures below 5-6 °C, while for *C. finmarchicus* and *C. helgolandicus* peak abundance is from 4 to 9 °C and from 13 to 17 °C, respectively (Bonnet et al. 2005; Carstensen et al. 2012; Hirche 1997; Sundby 2000). All four species contribute significantly to the biomass of their respective ecosystems, where they play the central role in the food web dynamics (Bonnet et al. 2005; Falk-Petersen et al. 2007, 2009; Hirche and Kosobokova 2007). However, they most likely support different food webs, due to differences in phenology (timing of reproduction) and energy-rich lipid content (Beaugrand et al. 2003; Conover 1988; Falk-Petersen et al. 2009).

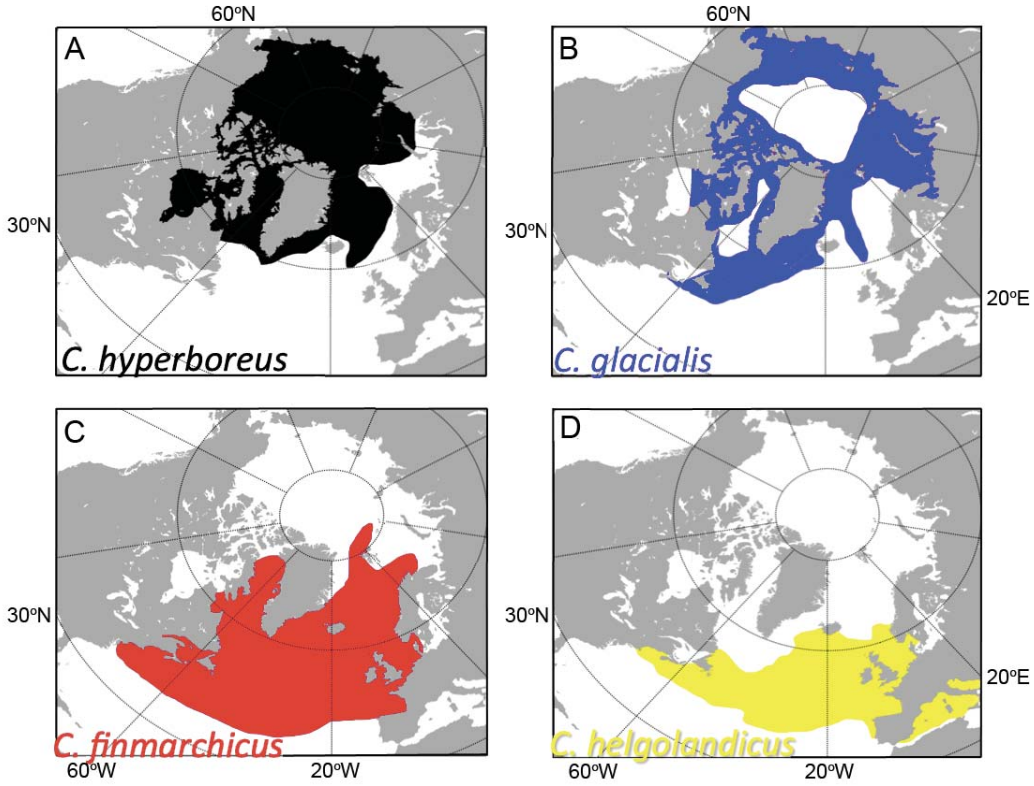

Figure 1. *Calanus* species distribution in the North Atlantic and the Arctic: (A) *C. hyperboreus*, (B) *C. glacialis*, (C) *C. finmarchicus*, (D) *C. helgolandicus*. Species distributions are based on data from Bonnet et al. (2005), Conover (1988), Falk-Pedersen et al. (2009), Fleminger and Hulsemann (1977), Sundby (2000) and Yebra et al. (2005).

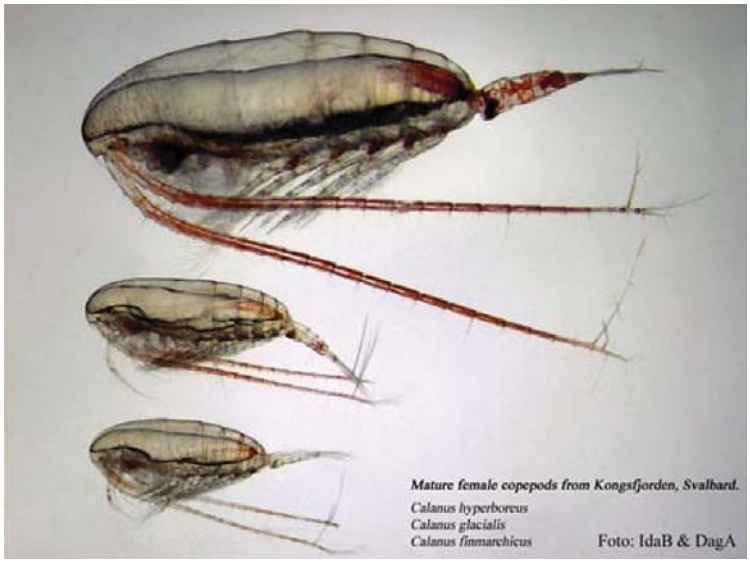

Figure 2. Appearance and size difference among three *Calanus* species: *C. hyperboreus* (top), *C. glacialis* and *C. finmarchicus* (bottom). Photo reprinted from Berge et al. (2011) with permission (License number 3675340618222).

Calanus species are generally very similar in morphology (Fig. 2) and life cycle, which includes eggs, six naupliar stages (NI-NVI) followed by five copepodite stages (CI-CV) and then maturation to the adult stage (male or female, CVIm/f) (Fig. 3). The annual life cycle of *C. finmarchicus* follows this pattern: eggs are laid by females either before or during the spring bloom, the eggs hatch and develop through naupliar and copepodite stages to pre-adult CV stage, when they descend to depth by early to late summer for overwintering in a resting state (diapause). In early spring, CVs migrate to the surface to moult to the adult stage and mate (reviewed by Head et al. 2013). The duration of life cycle is variable among *Calanus* species, as well as within each species living under different environmental conditions. Thus, *C. helgolandicus* typically has multiple generations per year (Bonnet et al. 2005), while *C. finmarchicus* has an annual life cycle throughout much of its range, but also can have multiple generations per year or one generation can exceed one year (Conover 1988; Heath et al. 2008; Hirche et al. 2001). *Calanus glacialis* has a life cycle of 1-3 years, while *C. hyperboreus´* cycle is 1-6 years (reviewed by Falk-Petersen et al. 2009). *Calanus hyperboreus* primarily spawns at depth, independent of food during winter and early spring (Hirche and Niehoff 1996; Niehoff et al. 2002), while *C. glacialis* and *C. finmarchicus* spawn near the surface around the time of the spring bloom (Madsen et al. 2001), due to their

dependence on food availability, especially the smaller and less lipid-rich *C. finmarchicus* (Niehoff and Hirche 1996).

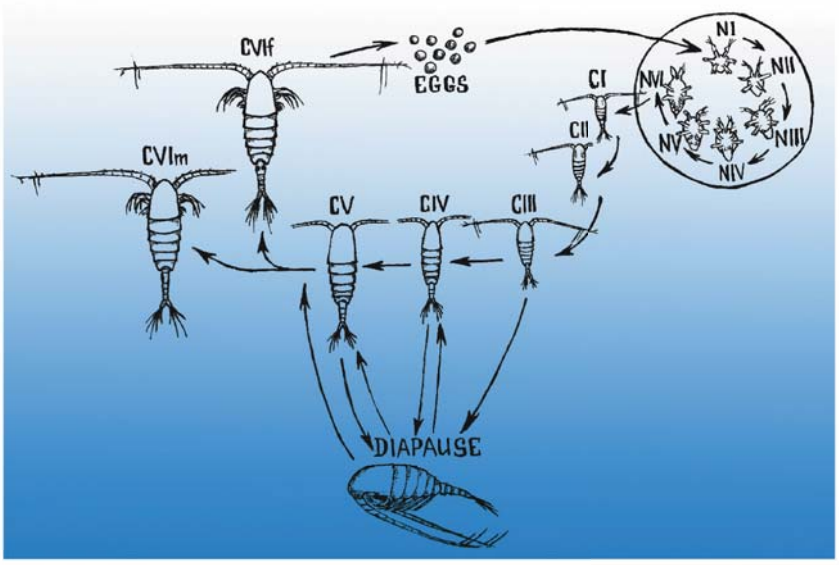

Figure 3. Life cycle in *Calanus* species. Eggs hatch and develop through six naupliar (NI-NVI) and five copepodite (CI-CV) stages and moult into male (CVIm) or female (CVIf). Copepodite stages V, IV, and III are usually the first stages that undergo diapause in *C. finmarchicus*, *C. glacialis*, and *C. hyperboreus*, respectively.

Despite the high abundance of *Calanus* species, their ecological importance, and the plethora of published studies, there is still limited knowledge and understanding at species, population and genome levels. Among these knowledge gaps, unbiased species identification, population genetic structure, tolerance and response to stressful conditions on the molecular level could be seen as the most important ones, particularly for prediction of climate change effects on *Calanus* species, and therefore, on functioning of ecosystems in the North Atlantic and the Arctic.

1.2. Gaps in knowledge about *Calanus*

1.2.1. Species identification

Correct species identification is the basis for any study, particularly for zooplankton, where cryptic species are common (Knowlton 1993, 2000). High morphological similarity of *Calanus* species (Fig. 2) – with subtle variations in secondary sex characteristics (Fleminger and Hulsemann 1977; Frost 1974) – present a persistent challenge for the identification of individuals to species level, particularly during early

developmental stages (Bucklin et al. 1995b; Lindeque et al. 1999). Currently, species identification is frequently based on individual body length and the geographical location of collection (e.g., Hirche et al. 1994; Kwasniewski et al. 2003; Undstad and Tande 1991). However the body length is temperature-dependent (Campbell et al. 2001; Wilson et al. 2015), and most likely leading to a greater overlap in body length between the sibling species than previously thought (Kwasniewski et al. 2003; Parent et al. 2011; Wilson et al. 2015). Thus, morphological species identification is not sufficient and may lead to consistent misidentification, particularly in regions where distributions of several *Calanus* spp. overlap, such as the North Atlantic, as has been shown with genetic methods (Gabrielsen et al. 2012; Lindeque et al. 2006; Parent et al. 2012).

Genetic methods for identification of *Calanus* species have followed the evolution of molecular technologies starting from allozymes (Sywula et al. 1993) and Sanger sequencing (Bucklin et al. 1995a; Hill et al. 2001) to microsatellites (Provan et al. 2012) and metabolomics (Hansen et al. 2013b). Nevertheless, all developed methods have potential drawbacks. For example, revolutionary in its time, species identification based on sequences of mitochondrial genes could result in unambiguously discrimination of most of *Calanus* species, including four species from the North Atlantic (Bucklin et al. 1992, 1995; Hill et al. 2001; Lindeque et al. 1999). However, the method may also lead to an erroneous identification of cross-species hybrids, due to the maternal inheritance of mitochondria (Parent et al. 2012). In contrast to mitochondrion-based methods, an application of 10 nuclear microsatellite loci allowed delimitation of *C. finmarchicus*, *C. glacialis* and their hybrids (Parent et al. 2012). However, the microsatellites markers were developed from *C. finmarchicus*, and cross-amplification of such microsatellites in *C. glacialis* may result in size homoplasy (when different alleles are identical in size but not identical by descent) and lead to biased conclusions about species identification (e.g., Curtu et al. 2004). Advances in metabolomics resulted in identification of several species-specific metabolites in *C. finmarchicus*, *C. glacialis*, and *C. hyperboreus* (Hansen et al. 2013b); however, the technique requires rather advanced equipment and is impossible to combine with

other applications, such as genetics. Thus, a reliable and routine method for *Calanus* species identification is still lacking, but is essential for resolving of the potential hybridization between *Calanus* species and further species-specific studies of physiology and genetics, distribution and reproductive ranges, responses to environmental factors and, consequently, potential impacts of climate change.

1.2.2. Response to climate change

Numerous species, communities and ecosystems worldwide are affected by recent global climate change (Beaugrand 2004; Edwards et al. 2013; Hickling et al. 2006; Parmesan 2006; Parmesan and Yohe 2003; Southward et al. 1995; Thackeray et al. 2008). Climate-induced changes are especially pronounced in polar and marine ecosystems (Gilg et al. 2012; Goberville et al. 2014; Hansen et al. 2006; Harris et al. 2014; Hoegh-Gulberg 2010; Perry et al. 2005) and are displayed as shifts in species distribution and phenology, which are consequently leading to changes in trophodynamics and ecosystem shifts (Beaugrand 2014; Beaugrand et al. 2002, 2003, 2015; Doney et al. 2012; Edwards and Richardson, 2004; Parmesan 2006; Richardson 2008). One of the most striking examples of such climate-driven changes is the northward shift of zooplankton assemblages in the North East Atlantic between 1960 and 1999, with the estimated speed of 260 km per decade (Beaugrand et al. 2002).

Calanus finmarchicus and *C. helgolandicus* are a part of this large-scale re-distribution of zooplankton in the North East Atlantic (Fig. 4A and 4B) (Beaugrand et al. 2002). For instance, in the North Sea *C. helgolandicus* substituted *C. finmarchicus* within 40 years, with an increase in abundance from 20 % to 80 % (Reid et al. 2003) and currently extends into the southern area of the Norwegian Sea (Ellertsen and Melle 2011). Such replacement of species with different lipid richness and seasonal cycles (Bonnet et al. 2005) has already had dramatic consequences for the recruitment of commercially important fish stocks such as cod (*Gadus morhua*) and salmon (*Salmo salar*) (Beaugrand and Reid 2003; Beaugrand et al. 2003; Olsen et al. 2011), as well as for the biological carbon pump (Beaugrand et al. 2010).

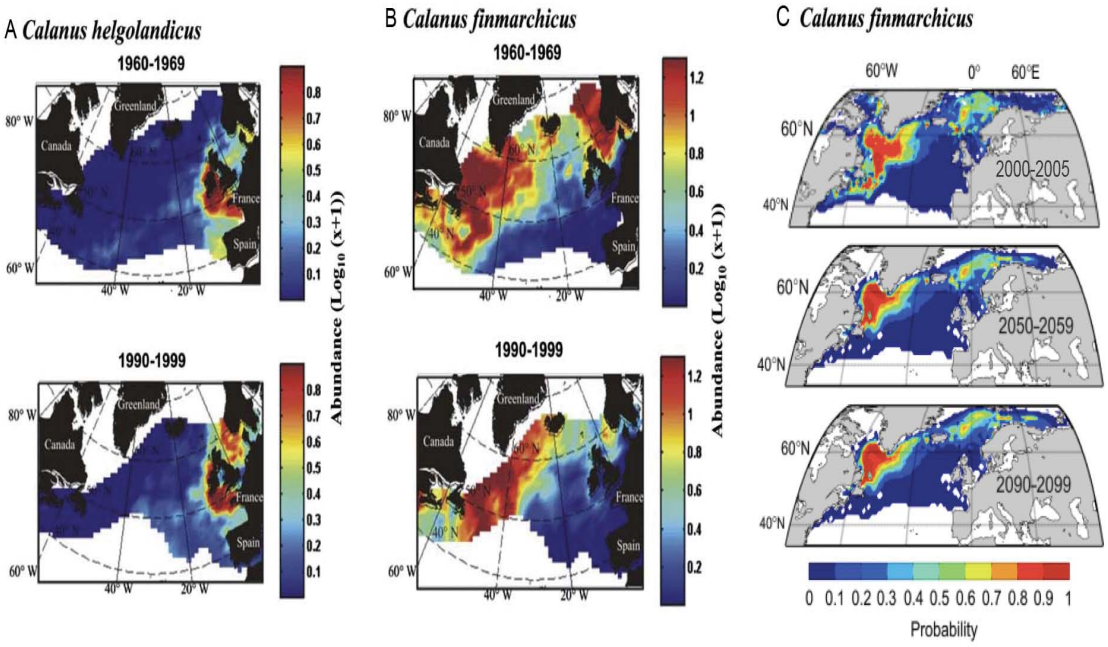

Figure 4. Distributional shifts in *Calanus helgolandicus* (A) and *Calanus finmarchicus* (B and C). Predicted distributions (C) based on moderate scenario B2 of future climate using nonparametric probabilistic ecological niche modelling (Reygondeau and Beaugrand 2011). Panels A and B were modified from Bonnet et al. (2005) with permission (License number 3675341020054); panel C was modified from Reygondeau and Beaugrand (2011) with permission (License number 3675341416082).

On the scale of North Atlantic, distributional changes have been detected since 1960 for all four *Calanus* species, with *C. finmarchicus* having a consistent poleward shift, *C. helgolandicus* expanding in all directions, and *Calanus glacialis* and *C. hyperboreus* showing a slight southward shift in the Scotian and Newfoundland shelf regions (Chust et al. 2013). This surprising southward shift of the Arctic species could be explained by the outflow of cold water from the Arctic (Chust et al. 2013); however, additional studies are needed to assess the species response throughout their main distributional centres in the Arctic (Fig. 1A and 1B). Thus, for example, in the Barents Sea, abundance of *C. glacialis* and *C. hyperboreus* declined after 2004 (Dalpadado et al. 2012). With continuing climate change, the distribution of *C. finmarchicus* is predicted to shift further into the Arctic, with a reduction in abundance (Fig. 4C) (Beaugrand et al. 2013; Helaouët et al. 2011; Reygondeau and Beaugrand 2011; Wassmann et al. 2011), while *C. glacialis* may be forced out of suitable habitat on continental shelves and only be able to maintain viable populations in cold Arctic fjords (Slagstad et al. 2011). Due to such re-distribution and expectations that – in terms of physiological tolerances –

Arctic warming will likely benefit *C. finmarchicus* more than *C. glacialis* (Kjellerup et al. 2012), significant changes in food-web dynamics and secondary production are expected, particularly regarding the current indigenous natural predators such as fish and/or plankton eating seals, whales and birds (Falk-Petersen et al. 2007; Whitfield 2008).

Water temperature appears to be the major driver of distributional shifts in *Calanus* species (Chust et al. 2014; Helaouët and Beaugrand 2009; Helaouët et al. 2011; Reygondeau and Beaugrand 2011), but the molecular basis of physiological responses to temperature range remains largely unknown. Organismal thermal tolerance windows have evolved to match surrounding temperature ranges, while minimizing physiological costs (Hofmann and Todgham 2010; Pörtner and Farrell 2008). When environmental temperatures shift closer to one edge of an organism's thermal window, the individual's physiological performance (e.g., growth, reproduction) will be negatively impacted (Pörtner and Farrell 2008; Somero 2012). Since a thermal window may be species- or even population-specific, the strength of an impact may vary, but usually organisms with narrow thermal windows or living close to their physiological limits will be most affected (Tomanek 2010).

Molecular responses to temperature are modulated by phenotypic plasticity (phenotypic/physiological acclimatization to changing conditions) and genetic adaptation (genetic evolution through natural selection to new environmental conditions) and may play important roles in persistence of the species and may change predicted distributional changes. Acclimatization via physiological plasticity could be reached rapidly – within one generation – for instance via differential regulation of gene expression (Ferea et al. 1999), particularly of heat shock proteins (HSPs), which act as molecular chaperones and support functioning of the organism under stress conditions (Feder and Hofmann 1999; Hofmann and Todgham 2010). Nevertheless, physiological plasticity (good performance in all environments) is limited by physiological/energetic costs and evolution (Sørensen et al. 2003).

For the long-term persistence of populations, evolutionary adaptation becomes essential. Evolutionary rates in marine plankton can be rapid (Dawson and Hamner

2005), but it is unclear whether the rate of adaptation will match the rate of climate change (Jump and Peñuelas 2005). In zooplankton, rapid evolutionary changes might be promoted by large population size, which can facilitate the appearance of beneficial mutations, short generation span, and high genetic diversity (Peijnenburg and Goetze 2013; Reusch 2013). In addition, patterns of gene flow may favour evolution: high dispersal can distribute genetic variants across the species distribution and expose them to different selection pressures (Norris 2000), or spatially-isolated populations may adapt to local conditions. In the latter case, species population are likely to respond differentially to global climate change (Goetze et al. 2011; Peijnenburg and Goetze 2013; Reusch and Wood 2007). Thus, studies of population structure and transcriptome-wide responses to temperature are required to increase our understanding of species performance and distribution in changing climate, and therefore, to predict "winning" and "losing" populations and/or species, and consequently, climate-related changes in the ecosystem diversity and structure, functioning and provision of services to society.

1.2.3. Population structure

Genetic differentiation in marine holoplankton is usually expected to be low due to large population sizes, wide geographic ranges and high potential for dispersal and, thus, for gene flow. However, in the last two decades, substantial genetic structure has been detected in many oceanic zooplankton species and other pelagic marine organisms across their geographic range (Andrews et al. 2014; Goetze 2003, 2005, 2011; Hauser and Carvalho 2008; Jorde et al. 2015; Peijnenburg et al. 2004, 2005). These findings highlight the influence of factors such as isolation by distance, oceanic currents, behavioural limits to dispersal, selection, or recent population history, which can promote population differentiation (Palumbi 1994). In addition the strength of genetic isolation may vary among very closely related species (Aarbakke et al. 2014; Chen and Hare 2011; Goetze 2005), and is likely linked to the ecological requirements of the species (Peijnenburg and Goetze 2013). In *Calanus,* for example, strong population differentiation was detected in *C. helgolandicus* (Papadopoulos et al. 2005;

Yebra et al. 2011) and *C. pacificus* (Nuwer et al. 2008) but no differentiation was found between populations of *C. sinicus* (Huang et al. 2014) (Table 1). However, population genetic structure studies in *Calanus* have mainly utilized sequence variation of one mitochondrial gene (Table 1). Since one or few markers may have insufficient power to resolve subtle population genetic structure, lack of differentiation between populations may be an artefact and should be reviewed with more markers.

Among *Calanus* species, the population structure of *C. finmarchicus* has been the most intensively investigated; however, the results of the different studies are rather conflicting (Table 1). An analysis with six microsatellites and cytochrome b sequence variation revealed no significant genetic differentiation (Provan et al. 2008), while studies with 16S rRNA sequence variation or several SNPs showed evidences for basin-scale differentiation between North East and North West Atlantic populations, which may reflect entrainment in ocean gyres (Bucklin and Kocher 1996; Bucklin et al. 1996; Unal and Bucklin 2010). Thus, it is still unclear if the significant variation in life history traits observed over the geographic distribution of *C. finmarchicus* (e.g., numbers of generations per year, timing of reproduction, seasonal patterns of abundance; Heath et al. 2000, 2004; Melle et al. 2014a,b; Planque et al 1997) is determined by genetic differentiation, but it is becoming a crucial question in the light of climate change (Wilson et al. 2015). For instance, locally adapted populations of one species will most probably have different responses to shared physical or climatic forcing (Goetze et al. 2011). Overall, development of genetic resources and markers throughout the genome is needed to resolve conflicting results in case of *C. finmarchicus* and obtain more detailed resolution of genetic structure in other *Calanus* species.

Table 1. Genetic structure studies in *Calanus* species.

Species	Reference	Marker	Area	Main conclusion
Calanus finmarchicus	Sywula et al. 1993	Allozymes, 2 loci	Svalbard fjord	No differentiation
	Bucklin et al. 1996	mtDNA haplotypes, 16S rRNA	North Atlantic	Meso- and large-scale differentiation between east and west North Atlantic and within west North Atlantic
	Bucklin and Kocher 1996	mtDNA haplotypes, 16S rRNA	North Atlantic	Large-scale differentiation between east and west North Atlantic
	Kann and Wisher 1996	Allozymes, 5 loci mtDNA haplotypes, 16S rRNA and cytochrome b	Gulf of Maine	No differentiation but evidence for subdivision near the area boundaries
	Bucklin et al. 2000a	nucDNA haplotypes, 3 loci	Icelandic waters	Differentiation between central North Atlantic and Arctic
	Bucklin et al. 2000b	mtDNA haplotypes, 16S rRNA	Norwegian fjords	No differentiation
	Provan et al. 2008	Microsatellites, 6 loci mtDNA haplotypes, cytochrome b	North Atlantic	No differentiation
	Unal and Bucklin 2010	nucDNA SNPs, 24 loci, 3 genes	North Atlantic	Meso- and large-scale differentiation reflecting pattern of oceanic gyres
Calanus helgolandicus	Bucklin et al. 2000b	mtDNA haplotypes, 16S rRNA	Norwegian fjords	No differentiation
	Papadopoulos et al. 2005; Unal et al. 2006	mtDNA haplotypes, cytochrome oxidase I	North East Atlantic	Large-scale differentiation between North East Atlantic and Adriatic Sea
	Yebra et al. 2011	mtDNA haplotypes, 16S rRNA	North East Atlantic and Mediterranean	Meso- and large-scale differentiation between and within sea basins

Table 1 continued

Species	Reference	Marker	Area	Main conclusion
Calanus glacialis	Nelson et al. 2009	mtDNA haplotypes, 16S rRNA	Arctic and North Pacific	Large-scale differentiation between Pacific and Arctic Ocean
Calanus pacificus	Bucklin and Lajeunesse 1994	mtDNA haplotypes, 16S rRNA	North Pacific	Differentiation into two subspecies
	Nuwer et al. 2008	mtDNA haplotypes, cytochrome oxidase I	North Pacific	Strong large- and meso-scale differentiation
Calanus sinicus	Kozol et al. 2012	mtDNA haplotypes, cytochrome oxidase I	Sea of Japan and North West Pacific	No differentiation
	Huang et al. 2014	mtDNA haplotypes, cytochrome oxidase I	Shelf along China coast	No differentiation
	Schizas et al. 2014	nucDNA haplotype, 1 loci mtDNA haplotypes, cytochrome b	North West Pacific	Contradiction: meso-scale differentiation with nucDNA and no differentiation with mtDNA
Calanus agulhensis	Kozol et al. 2012	mtDNA haplotypes, cytochrome oxidase I nucDNA haplopytes, 2 loci	Along coast of South Africa	No differentiation

Key: mtDNA, mitochondrial DNA; nucDNA, nuclear DNA.

1.2.4. Genomic resources

Although copepods are among the most numerous and important species in aquatic ecosystems, their genomic resources are still limited, particularly when compared to insects (Bron et al. 2011). There is still no publically available nuclear genome for any copepod, and the most closely related species with a reference genome is the highly-derived cladoceran, *Daphnia pulex* (Colbourne et al. 2011). The major obstacle to genome sequencing in copepods, including *Calanus* spp., is their very large genome size, with a range of C- values (i.e., amount of DNA contained within a haploid nucleus) from 6.48 pg (≈ 6.34 Gb) for *C. finmarchicus* to 12.46 pg (≈12.19 Gb) for *C. hyperboreus* (McLaren et al. 1988). Even mitochondrial genomes of *Calanus* spp. are difficult for traditional Sanger sequencing with long PCR amplification, as the published mitochondrial genomes of *C. sinicus* (Wang et al. 2011) and *C. hyperboreus* (Kim et al. 2013) show multiple long non-coding regions and large-scale gene rearrangements, in contrast to the usually conserved gene order among vertebrates (Bron et al. 2011).

Sequencing of the transcribed part of genome is often the easiest solution to reduce the complexity of the genome. *Calanus finmarchicus* is one of the leading species among copepods for the amount of available transcriptome data, together with the parasitic copepods of salmon *Lepeophtheirus salmonis* and *Caligus rogercresseyi*, and a model species in evolutionary genetics and ecotoxicology *Tigriopus californicus* (Bron et al. 2011). A large number of Expressed Sequence Tags (ESTs) was sequenced in *C. finmarchicus* to investigate diapause changes (Tarrant et al. 2008), physiological responses to environmental variations (Lenz et al. 2012) and a mixture of environmental stressors (Hansen et al. 2007), resulting in 11,859 ESTs (≈ 7.55 Mb) (http://www.ncbi.nlm.nih.gov/gquery/?term=Calanus, accessed on May 5 2015). The advances in next generation sequencing have brought new opportunities with sequencing of transcriptomes of *C. finmarchicus* (Lenz et al. 2014; Tarrant et al. 2014), *C. sinicus* (Ning et al. 2013; Yang et al. 2014) and *C. glacialis* (Ramos et al. 2015) resulting in 60,000 - 100,000 unique transcripts for each species.

For other *Calanus* species, genomic resources are even scarcer, being limited to a handful of mitochondrial (16S rRNA, COI, Cytb) and nuclear genes (citrate synthase,

28S rRNA, ITS-1) (Bucklin etal 1995; Hill et al. 2001; Kozol et al. 2012; Papadopoulos et al. 2005; Schizas et al. 2014; Unal et al. 2006; Yabra et al. 2011). Access to more complete genome and transcriptome resources for *Calanus* spp. would facilitate research in ecological and applied, genetic and evolutionary fields, consequently leading to better understanding of *Calanus* potential to cope with climate change.

1.3. Next generation sequencing tools in molecular ecology

Over the past decade, sequencing technologies have undergone significant changes toward high-throughput parallel sequencing, and research has shifted away from an application of automated Sanger sequencing (considered as "first-generation" technology) to newer methods, next-generation sequencing (NGS) technologies (Metzker 2010). Different NGS platforms have appeared in the marketplace since 2005: 454 pyrosequencing (Roche), Genome Analyzer (Illumina), SOLiD (Applied Biosystems), nanopore (NanoPore), ion semiconductor (Applied Biosystems) and some others. Most of these differ in the method of template preparation and sequencing, read-length, run time, and data throughput, but all lead to inexpensive production of large volumes of sequence data (Metzker 2010). Since the first application of NGS for a species without previous genomic data, the wasp *Polistes metricus*, (Toth et al. 2007), NGS technologies have been revolutionizing the fields of ecology, genetics, evolution, and conservation (Ekblom and Galindo 2011; Rokas and Abbot 2009).

Continuous development of NGS technologies is leading to increases in throughput and length of sequenced reads, and decreases in error rate and sequencing prices. The flip-side is that the amount of data produced has increased dramatically (e.g., one run on HiSeq 2500 instrument generates up to 1 TB of data), requiring massive data storage facility and new bioinformatics solutions to effectively analyze the sequence data. Bioinformatics rather than actual sequencing is now the bottleneck (Funari and Canosa 2014; Stillman and Armstrong 2015), demanding time and qualified molecular ecologists to make biological sense of the genomics data (Ekblom and Galindo 2011; Rossetto and Henry 2014; Tautz et al. 2010). The complexity and amount of data are challenging, particularly when development of bioinformatic tools is lagging behind

application of new NGS technologies and methods. Nevertheless, with some time lag, new bioinformatics algorithms tailored to various NGS applications are being developed (e.g., open source software tools for bioinformatics available at http://www.bioconductor.org (Huber et al. 2015) or updated list of software at http://seqanswers.com/wiki/Software).

A multitude of new NGS approaches and tools makes it possible to investigate genome, transcriptome and gene expression, as well as to develop new markers and assess population diversity and differentiation in non-model species (Fig. 5). Thus, questions in phylogeny and phylogeography (McCormack et al. 2013), population and ecological genomics (De Wit et al. 2012; Stapley et al. 2010), adaptation and selection (Narum et al. 2013; Schoville et al. 2012), and, consequently, climate change responses can be addressed directly in ecologically important species.

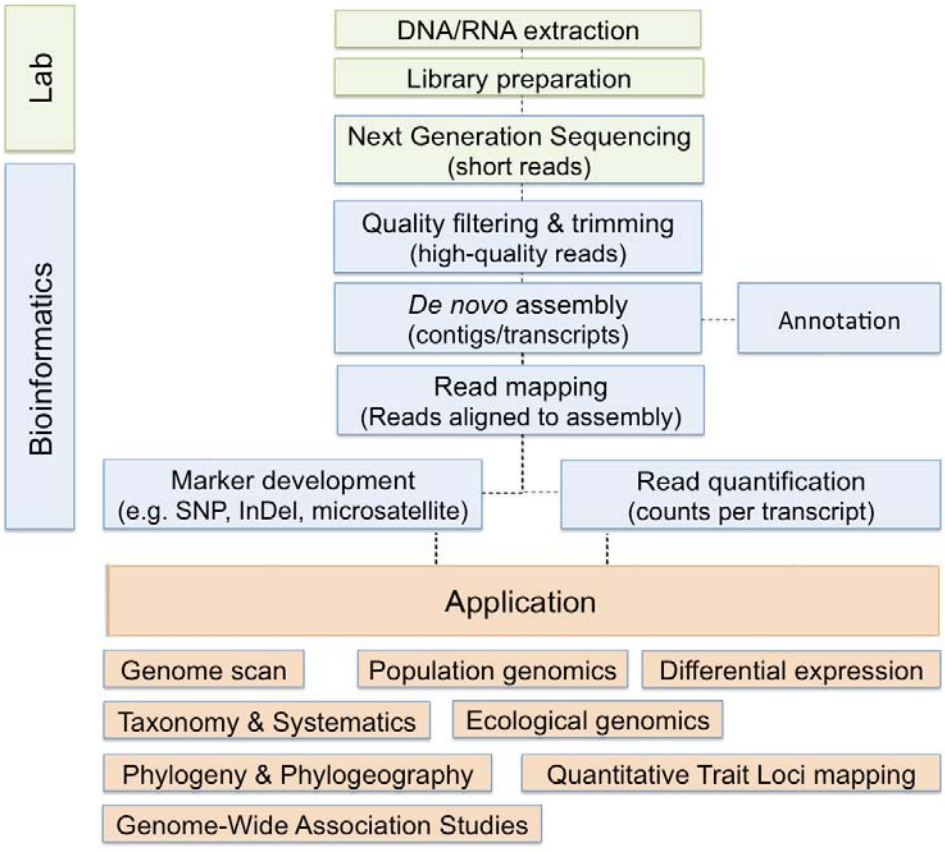

Figure 5. Flow chart of a typical NGS experiment in non-model species (without reference genome/transcriptome) and possible NGS applications.

1.3.1 Genome and transcriptome sequencing

Progress in NGS and development of bioinformatic software to assemble short NGS reads have resulted in the genome or transcriptome of many species. Thus, after launching in 2008 “The 1000 Genomes Project” (http://www.1000genomes.org), devoted to sequencing and characterising genetic variation in 1000 humans, similar projects have been started and included sequencing of 1000 transcriptomes of plants (“1KP”, https://sites.google.com/a/ualberta.ca/onekp/), more than 10,000 genomes of vertebrates (“Genome 10K”, http://genome10k.soe.ucsc.edu) and 5,000 genomes of arthropods (“i5k”, http://www.arthropodgenomes.org/wiki/i5K). Nevertheless, complete genome/transcriptome sequencing and assembly still requires costs, expertise and infrastructure that are beyond reach for most molecular ecology research groups. Thus, most of the NGS applications in non-model species rely on partial *de novo* assemblies that are performed from sequencing reads prior to downstream applications, such as marker development (e.g., Cook et al. 2011, Karam et al. 2014), comparative genomics (e.g., Künstner et al. 2010), population differentiation (e.g., Chu et al. 2014) and local adaptation (De Wit and Palumbi 2013) (Fig. 5).

In the case of species with a large genome, the transcriptome could be more accessible, since it is generally much smaller than the corresponding genome (Riesgo et al. 2012). Nevertheless, the analysis leading to *de novo* assembly of genomes and transcriptomes requires similar steps (Fig. 5), once mRNA is reversely transcribed into complementary DNA (cDNA) during library preparation. First, the short reads produced by NGS platforms should be quality controlled: trimming of any adapter sequences and bases that are below a quality threshold, and filtering reads if read length is too small. High-quality short reads are then assembled by overlap of similar regions into longer sequences (contigs), which represent transcripts or parts of a genome.

De novo assembly is still a difficult and computationally demanding task, particularly in case of large amounts of short reads, such as are produced by NGS. The most memory efficient solution is an algorithm that finds the Eulerian path through a de Bruijn graph (Pevzner et al. 2001). It breaks reads in pieces of defined length (k-mers),

merges and counts identical overlaps of k-mers, while different k-mers result in a graph branching or bubbles that represent genetic variation or a sequencing error. At the end, the path through the graph outlines a draft assembly (Schliesky et al 2012). The genome graph can be resolved using the degree of coverage for each k-mer, since the coverage should be the same across genome. However, a broad range of transcript abundance in a cell results in highly uneven coverage across the transcriptome assembly graph, and thus, requires an application of transcriptome-specific assemblers (Schliesky et al. 2012). Nevertheless, some reads may be incorrectly discarded as mistakes or repeats, or joined up in the wrong places or orientations, while sequences of similar paralogues may be collapsed together (Baker 2012; Konczal et al. 2014).

Obtained *de novo* genome/transcriptome assembly should reflect the real genome/transcriptome, however, since the true reference is unavailable, the evaluation of assembly by comparison is impossible. Nevertheless, accuracy and completeness of *de novo* assemblies should be assessed, with metrics such as: number of assembled contigs, proportion of reads that were assembled, length of contigs, length of the N50 contig (smallest contig above which 50% of an assembly would be represented), and average percentage of protein sequences discovered (Baker 2012; Schliesky et al. 2012). Since a high-quality and complete assembly is a prerequisite for non-biased downstream analysis (e.g., Singhal 2013), the effectiveness of *de novo* assemblers for genome and transcriptome has been widely compared and discussed (e.g., Gurevich et al. 2013; Singhal 2013; Zhang et al. 2011).

To fully benefit from the assembled reference genome/transcriptome, reconstructed contigs should be functionally annotated (i.e., assigned to known protein function). Annotation can be performed by comparing contig sequences to known sequences in protein databases at NCBI (http://www.ncbi.nlm.nih.gov/) using a Basic Local Alignment Search Tool search (BLAST). Annotation success varies between different species and can be low, particularly for non-model species, due to unknown genes or large sequence divergence (Angeloni et al. 2012). Nevertheless, the procedure provides important knowledge about genes in non-model species, resulting,

for instance, in finding candidate genes for selection (Hohenlohe et al. 2010) or molecular pathways involved in response to stress (e.g., Schoville et al. 2012).

1.3.2 RNA-seq

In the NGS era, gene expression studies are not anymore limited to small-scale quantitative PCR analyses of candidate genes or defined design of microarrays (Wolf 2013). Instead, whole-transcriptome shotgun sequencing (RNA-seq) or high-throughput sequencing-based expression profiling of RNA (cDNA), can characterize the transcriptome, the sequences and abundances of transcripts from any sample. This approach can be used to characterise a species´ transcriptome and develop markers (Simon et al. 2009), as well as to profile differential gene expression between life stages (e.g., Lenz et al. 2014, Pérez-Porro et al. 2013), populations (Ji et al. 2013; Schoville et al. 2012), species (e.g., Gao et al. 2011; Wolf et al. 2010), or in response to stress (Liu et al. 2014). RNA-seq technologies, bioinformatics, analytical and statistical tools are still developing and require careful consideration. Nevertheless, RNA-seq validation with qPCR and microarray has confirmed the suitability of the approach (e.g., Kristiansson et al. 2009; Liu et al. 2014) and leads to encouragement of RNA-seq studies in non-model organisms (e.g., De Wit et al. 2012; Singhal 2013; Wolf 2013).

The rationale behind RNA-seq is that the number of reads mapped to a transcript is proportional to the transcript abundance in the original sample (t Hoen et al. 2008), thus providing a digital measure of transcript abundance. For species without a reference genome/transcriptome, reads can be mapped to the *de novo* assembled transcriptome (Fig. 5), or to the reference genome/transcriptome of a closely related species. Correct mapping of short reads can be a complicated problem due to sequencing errors and polymorphism, but it is important for accurate gene expression assessment (Wolf 2013). Although the number of mapped reads per transcript is used as a digital measure of transcript abundance, such correlation will be confounded by transcripts with different lengths (longer transcripts will have more reads than an equally expressed shorter one), by different sequencing effort between samples (more sequencing will result in more reads and higher transcript expression), and by a carry-

over effect of gene expression from a few genes to others (very high expression of one or few genes leads to reduction of available reads to others genes, resulting in artificially lower digital expression) (Wolf 2013). Normalization should be applied to correct for these issues, and currently different normalization techniques are implemented in the bioinformatics packages for differential expression analysis (e.g., in DESeq2 (Love et al. 2014); edgeR (Robinson et al. 2010)). An alternative is to use standardized spike-in RNA controls that are synthetic RNAs of known concentration, defined length and GC content, and can help to assure comparability across samples, protocols and platforms (Jiang et al. 2011).

Differential gene expression analysis, the statistical comparison of expression levels between transcripts and samples across treatment groups, is based on a statistical distribution (overdispersed Poisson or a negative binomial distribution) that captures the essential information about RNA-seq read count data (Young et al. 2012). Most software packages rely on such distributions and perform differential expression analysis with different levels of sophistication (for review see Rapaport et al. 2013; Soneson and Delorenzi 2013). Usually, differential expression analyses result in a set of candidate genes that differ between treatments or populations. However, for biologically meaningful interpretation, differentially expressed genes should be investigated for their individual function or for over-representation in some metabolic pathways using, for example, enrichment analysis or mapping the genes of interest directly to candidate metabolic pathways (e.g. KEGG (Ogata et al. 1999), Gene Ontology (GO) (Ashburner et al. 2000)).

1.3.3 Genetic marker development

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 2004). Historically, development of markers was difficult and expensive

for non-model organisms. However, the advent of NGS has revolutionized this by allowing the genome-wide markers in any organism and for low costs (Ekblom and Galindo 2011). Marker development is the most popular application of NGS. Thus, large number of genome-wide markers can be generated rapidly and cost-effective using whole genome shotgun sequencing, RNA-seq, or sequencing of reduced representation libraries (Davey et al. 2011). Markers dispersed throughout genomes enable more accurate estimations in population genetic studies, detection of adaptation and selection, genome-wide association studies (GWAS), QTL and linkage disequilibrium (LD) mapping projects, kinship and introgression assignments (Jakobsson et al. 2008; Novembre et al. 2008; Santure et al. 2010; Slate et al. 2009). Currently, these applications utilize microsatellites, InDels, and SNPs as markers of choice.

The type of marker and its final purpose will influence the standard workflow for NGS marker development (Fig. 5). For example, studies of selection and adaptive differentiation in natural populations require markers being located in genic regions of the genome. Thus, data mining from transcriptome sequencing will be the most rational (Imelfort et al. 2009). Depending on the purpose of study, sequencing could be performed for one or few individuals of same or different species (e.g., microsatellite or InDel discovery), in other cases, such as SNP discovery, many individuals from different populations of one species are required. Nevertheless, one sequencing project may result in parallel development of several types of markers (e.g., Choi et al. 2013).

The success of marker mining from a new sequencing project or archived data largely depends on advances in development of bioinformatics tools (Ruperao and Edwards 2014). Thus, software packages for detection of microsatellites from NGS data, such as msatcommander (Faircloth 2008) and MSatFinder (Thurston and Field 2005), allow for rapid screening of all contigs and typically result in discovery of several thousand microsatellite loci in a single individual (e.g., Hunter and Hart 2013). SNPs and InDels are usually detected simultaneously as single nucleotide variants (SNVs) using many individuals sequenced at sufficient depth. The most difficult problem in

detecting true SNPs is that sequencing and alignment errors will show a signal very similar to low-frequency SNP alleles (Ekblom and Galindo 2011). To reduce false positive SNPs, each step of the workflow (Fig. 5) should have a strict quality control, and appropriate SNP calling software should be selected (Singhal 2013). Tools for SNP calling are continuously evolving toward error-free SNP calling and genotyping, but special care should be taken in case of extreme read depth and pooled samples (Schlötterer et al. 2014).

Large-scale studies with genotyping of many individuals for the same loci, such as population differentiation, parental assignment or selection, may benefit from development of a marker panel. For example, thousands of SNPs that were developed from large-scale DNA or RNA sequencing could be arranged on a SNP-chip and routinely and quickly used for screening of hundreds of individuals (Davey et al. 2011). One of the caveats with this approach is that if markers were discovered on the basis of small population samples, genotyping of new populations will be biased towards alleles present in the original survey, and will result in ascertainment bias to high frequency variants (Helyar et al. 2011). Thus, ascertainment bias may significantly bias downstream genetic analyses, particularly of any statistical measure that relies on allele frequency (Albrechtsen et al. 2010). Ascertainment panel bias can be minimized by careful design of the marker panel, using data from multiple individuals from populations across the distributional range (Helyar et al. 2011). Alternatively, the bias can be avoided by application of genotyping-by-sequencing that allows simultaneous discovery of markers and genotyping (Davey et al. 2011).

1.3.4 Genotyping-by-sequencing

Although simultaneous discovery and genotyping of genome-wide variation has become feasible for tens of individuals with small genome sizes (< 500 Mb), 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 next-generation technology of a targeted fraction of the genome via various reduced-representation protocols. 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 and RAD-seq (reviewed by Davey et al. 2011). Each method has its strengths and weaknesses compared to others (Table 2) (Godden et al. 2012; McCormack et al. 2013; Schlötterer et al. 2014), and the best approach will depend on the research goal, the number and quality of samples to be assayed, and the available funding.

RNA-seq (section 1.3.2) can be viewed not only as differential expression method but also as method of genomic complexity reduction for GBS where only transcribed (coding) fraction of genome will be sequenced (Fig. 6A). Since mRNA isolation and sequencing does not require species-specific reagents, GBS with RNA-seq can be a cost-effective option for species with large genome and lacking a reference. However, in case of RNA-seq, uncertainty about population allele frequency estimation may arise from unequal levels of allele-specific gene expression, leading to mis-calling a heterozygous genotype as homologous (Schlötterer et al. 2014).

Amplicon sequencing (Ampli-seq, also known as parallel tagged sequencing) entails sequencing of selected PCR tags (Fig. 6B). Multiple loci of an individual are amplified and barcoded, pooled with tags of other individuals and high-throughput sequenced (Meyer et al. 2008). Ampli-seq may be the most cost-effective method for small- to medium-sized projects with few loci to amplify across individuals, such as phylogenetic analysis (Chan et al. 2010), environmental sampling and metagenomics (Fierer et al. 2008). However, it is less effective for investigation of genome-wide variation in hundreds of individuals, which is necessary for population structure studies.

Reduction of genome representation by Cap-seq (capture sequencing, also known as targeted resequencing or target enrichment), involves selective capture of defined genomic regions prior to NGS (Mamanova et al. 2009). The Cap-seq protocol (Fig. 6C) includes fragmentation of DNA and its hybridization with DNA or RNA probes (baits) either on an array or in solution and following sequencing of targeted DNA using NGS (reviewed by McCormack et al. 2013). Depending on the goals, target regions can be coding or non-coding, but the sequences of the target regions must be known *a priori* to design capture baits. Thus, Cap-seq methods can be applied in a medium-scale genomics studies with semi-model species, and can benefit from already fragmented DNA (e.g., formalin-fixed samples) and sample multiplexing prior to enrichment (Kenny et al. 2011).

RAD-seq (restriction-site associated DNA sequencing) uses restriction enzymes (endonucleases) to randomly sample the genome at locations adjacent to restriction enzyme recognition sites. In the original RAD-seq protocol (Baird et al. 2008), total DNA is digested with one restriction enzyme that cleaves DNA within its recognition sequences and then adaptors with the sequence primers are ligated to the restriction ends. Next, DNA is randomly fragmented, and only the fragments containing the adaptors are sequenced, resulting in selected sequencing of the short flanking regions of restriction sites (Fig. 6D). The number of restriction sites, and, therefore, flanking regions depends on the choice of restriction enzyme and species genome, allowing for flexibility to target different number of loci (Baird et al. 2008; McCormack et al. 2013). Further development of the method resulted in several modifications and improvements: 2bRAD (Wang et al. 2012), which uses a special enzyme cleaving DNA upstream and downstream of the restriction site; ezRAD (Toonen et al. 2013), which is similar to the original RAD protocol, but includes size-selection of fragments instead of fragmentation; and ddRAD (Peterson et al. 2012), which relies on digestion with two enzymes and consequent size-selection of the fragments. The latter is of particular interest, since it allows for flexible control over the number of sequenced fragments via selection of only those fragments that are flanked by both restriction enzymes and of a specific size range (Fig. 6E) (Peterson et al. 2012). Such strict control over fragment

number is beneficial for species with large genomes, where the number of recognition sites for even rare-cutting restriction enzymes could be too large to allow for effective coverage for NGS. Despite the great potential of RAD-seq protocols for GBS (Table 2) (Narum et al. 2013), it harbours a major drawback – possible SNP polymorphism within the recognition site – that may result in the loss of a fragment and consequent allele dropout from the genetic analysis and biased estimation of population parameters (Arnold et al. 2013; Gautier et al. 2013b).

Despite the advances in GBS and new opportunities for research, GBS, similar to any emerging method, remains challenging. Nevertheless, application of GBS is leading to a wide range of studies covering population genomics, association and QTL mapping, and development of genomic and SNP resources.

Table 2. Comparison between reduced-representation protocols for genotyping-by-sequencing.

Method	Method-specific bias	Involved costs	Advantage	Weakness
RNA-seq	Allele-specific gene expression (Result: biased allele frequency estimation)	Moderate to high	No species-specific reagents	RNA degradation
Ampli-seq	Primer specificity (Results: loss of loci, chimeric sequence)	Low to moderate	High on-target efficiency	PCR for each individual at each locus
Cap-seq	Sequence divergence between probe and target genome (Result: loss of locus)	High	Even coverage across contigs; Fragmented DNA could be used	Need in genomic resources; Custom design for each species
RAD-seq	Polymorphisms in recognition site (Result: null alleles)	Low	Flexibility in number of fragments	Unknown number of fragments if genome is unknown

Table is based on reviews by Davey et al. (2011), Harvey et al. (2013) and McCormack et al. (2013).

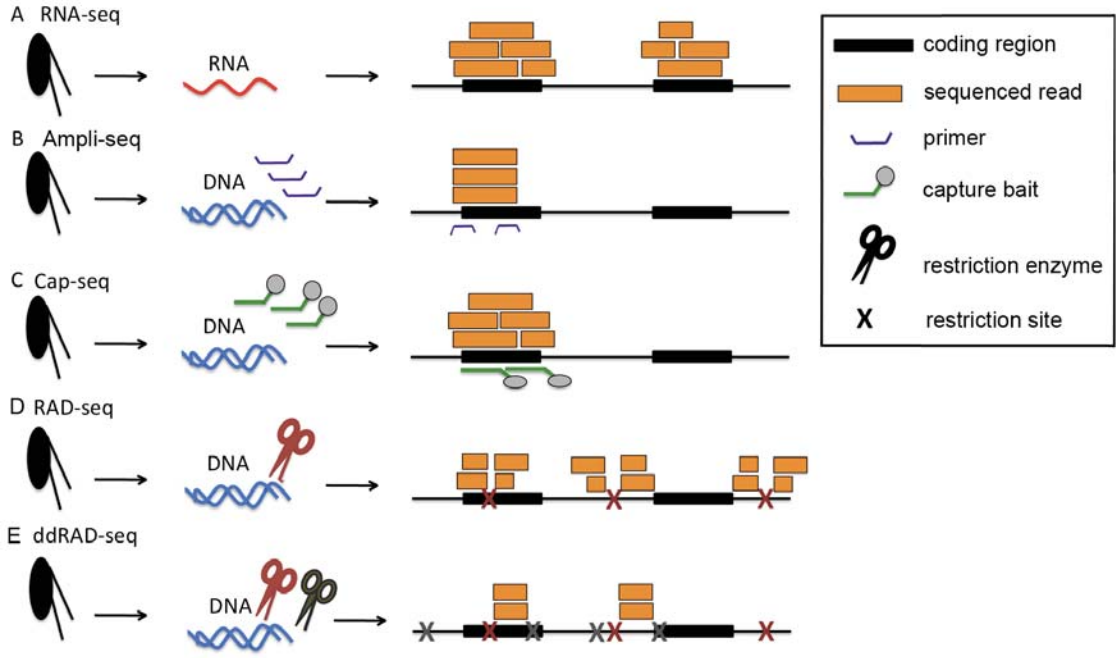

Figure 6. Reduced-representation strategies for genotyping-by-sequencing: A, RNA-seq (whole-transcriptome shotgun sequencing); B, Ampli-seq (amplicon sequencing); C, Cap-seq (capture sequencing); D, RAD-seq (restriction-site associated DNA sequencing); E, ddRAD-seq (double digest RAD-seq). For more details see text 1.3.4.

1.3.5 Pool-seq

Although sequencing costs are decreasing, many NGS applications require multiplexing of many individuals in one sequencing run, and therefore, increasing costs for preparation of multiple libraries. Alternatively, to reduce involved costs and time, multiplexing could be performed before library preparation. Such sequencing of pooled DNA/RNA samples from several individuals without unique barcodes that belong to a homogeneous group (population or treatment), is called Pool-seq (Fig. 7) (Futschik and Schlötterer 2010; Schlötterer et al. 2014).

The main limitation of Pool-seq is inability to assign reads from the pool to individuals. Thus, unequal DNA/RNA pooling or uneven amplification of DNA during library preparation can result in uncertain numbers of individuals or gene copies being represented among the sequencing reads for any location in the genome (Anderson et al. 2014; Ferretti et al. 2013; Gautier et al. 2013a). This uncertainty may influence estimation of population parameters and depends on the number of individuals

merged in the sequenced pool, the sequencing coverage of the pool, and the possibility of unequal contributions of each individual DNA to the final set of sequence reads (Gautier et al. 2013a). Several authors have developed a statistical foundation for the analysis of pooled samples and have provided recommendations for pool size, read depth, and bioinformatics tools to yield reliable estimation of population parameters and allele frequencies (Futschik and Schlötterer 2010; Gautier et al. 2013a; Schlötterer et al. 2014).

Although Pool-seq might not be optimal for studies that require genotype calling for individuals or data on linkage disequilibrium (e.g., Konczal et al. 2014), the approach is very powerful and cost- and resources-effective for transcriptome profiling (e.g., Kendziorski et al. 2005), marker development (e.g., Karam et al. 2014), and a broad range of genome-wide analyses. These include, for instance, an investigation of genome-wide patterns of differentiation and local adaptation in natural populations (e.g., Campana et al. 2015; Fabian et al. 2012; Karlsen et al. 2013) and experimental evolution (Tobler et al. 2014), discovery of candidate genes for tolerance (Turner et al. 2010), and detection of selective sweeps (Clément et al. 2013; Nolte et al. 2013).

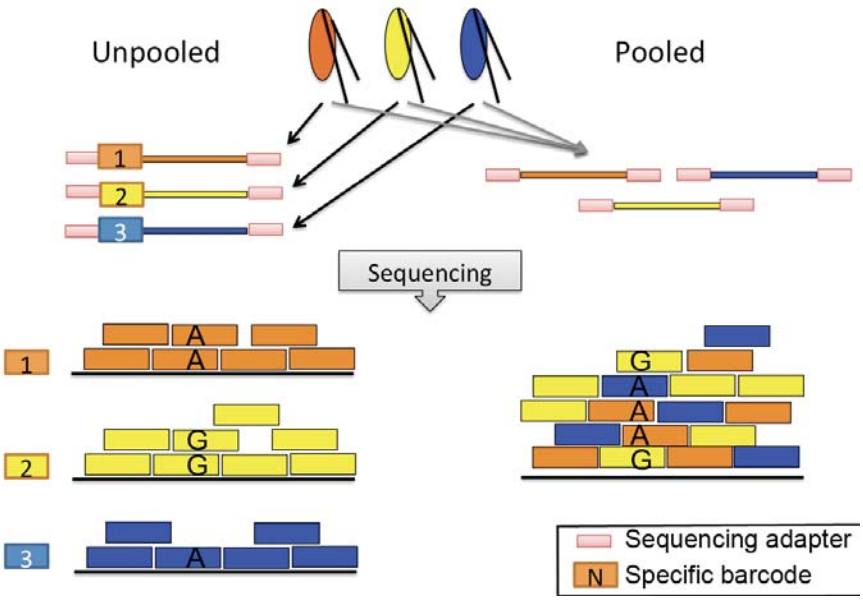

Figure 7. Individual-based and pooled-based sequencing.

Altogether, advances in NGS and bioinformatics open up new possibilities to answer questions that were impossible to address only a few years ago and to move from genetically model species toward studies of natural populations of ecologically important species or whole communities (Ekblom and Galindo 2011; Narum et al. 2013). Such development of ecologically key species into models for genomics would result in greatly enhanced ability to investigate patterns and processes at the molecular level that are relevant to the ecosystem functioning, particularly under climate change.

2. Objectives

The main aim of this dissertation is to address crucial knowledge gaps in *Calanus* species in the North Atlantic at species, population and genome levels using next generation sequencing tools, and thus, contribute to development of *Calanus finmarchicus* as model species for ecological and evolutionary genomics.

To address this aim, the project highlighted three objectives:

1. To establish robust and easy *Calanus* species identification in the North Atlantic (Papers I and II[*])

2. To assess the transcriptomic response to thermal stress in closely related species, *Calanus finmarchicus* and *C. glacialis*, with different thermal preferences (Paper III).

3. To investigate population connectivity and genetic structure of *Calanus finmarchicus*, a key zooplankton species in the North Atlantic (Paper IV).

[*] Paper II is an additional contribution to the three required papers for this dissertation, due to the fact that the paper has already been used as a chapter for PhD thesis by one of the co-authors. Nevertheless, work performed by the mentioned co-author and the author of this dissertation was independent, not overlapping, and of sufficient contribution. In addition, the paper is tightly connected to the dissertation and contributes to more complete presentation of the achieved results during PhD period of the author.

3. General discussion

3.1. Main findings

Since it is difficult and challenging to study an entire ecosystem with many elements and interactions among them, researchers often focus on the keystone species (Drinkwater et al. 2010). In this dissertation, I focus on the planktonic copepod *Calanus finmarchicus*, a key species for North Atlantic/Arctic marine ecosystems (reviewed in section 1), and address several knowledge gaps about the species (reviewed in section 1.2) using state-of-the-art next generation sequencing methods. The main findings of the research are:

- ***Calanus* species identification**: A panel of 12 nuclear insertion/deletion (InDel) markers was developed based on both genomic and transcriptomic sequences of *C. finmarchicus* and *C. glacialis* (Paper I). These new markers allow easy, fast and robust identification of all *Calanus* species present in the North Atlantic/Arctic Oceans. Their discriminatory power was much stronger than the microsatellites, allowing accurate detection of hybrids (Paper I and II). The markers were also used to validate a new morphological criterion – redness – to identify live females of *C. finmarchicus* and *C. glacialis* (Paper II).

- **Transcriptomic responses to thermal stress in *C. finmarchicus* and *C. glacialis***: Transcriptome-wide responses to realistic thermal stresses (up to 15 °C) for shorter (4 hours) and longer (6 days) exposures were investigated using RNA-seq and qPCR. Responses differed significantly between the two species, with marked up-regulation of genes in the boreal species (*C. finmarchicus*), in contrast to low-magnitude changes in gene expression in the Arctic species (*C. glacialis*) (Paper III). Together with other evidence, these results indicate that the Arctic species is likely to be more vulnerable to climate change in the Arctic ecosystem.

- **Population genetic structure of *C. finmarchicus*:** A novel approach for large-scale genome-wide genotyping of copepods, ddRAD-Pool-seq, was applied and evaluated for investigation of population genetic structure in *C. finmarchicus*. This approach was challenging due to the large genome of *C. finmarchicus*, nonetheless this method outperformed traditional microsatellite markers in providing evidence of both within- and between-basin differentiation in the North Atlantic (Paper IV).

- All together, the work in this dissertation has generated substantial genomic resources for several ecologically important species of *Calanus* (Paper I, III, IV), particularly unique genome sequences for *C. finmarchicus* and *C. glacialis*.

This dissertation thus contributes to filling several knowledge gaps, including species identification of *Calanus* in the North Atlantic Ocean, and transcriptomic responses to thermal stress. However some questions were only partially answered, thus, population genetic structure of *C. finmarchicus* still requires more investigation on finer spatial and temporal scales. In addition, new questions have been raised, including the existence and frequency of hybrids; basis of the weak transcriptomic stress response in the Arctic *C. glacialis* and the consequences for the species under climate change. Overall, all these findings contribute to a new view of *C. finmarchicus* as a model species for ecological and evolutionary genetics, thus fulfilling overall primary aim of the dissertation. In addition, the research performed within the dissertation may serve as a guide for similar studies in other copepod taxa, and contribute to studies on effects of climate change on species of *Calanus* and their corresponding ecosystems.

3.2. *Calanus finmarchicus* as a model for ecological genetics

Ecosystem properties and function results from interactions between species and their critical ecological traits. Thus, phenotypic variation in such traits may lead to a

cascade of effects, from change within a single species, to alterations of ecological interactions and community composition, and consequently to ecosystem changes (Miner et al. 2012). Since trait expression is the combination of environment and individual genotype, identification and investigation of key genes and pathways underlying ecologically important traits is an important task in ecological genetics, particularly in light of global climate change (Reusch and Wood 2007). Earlier, connection between phenotypes and underlying genotypes was possible only for model species, which are often limited to artificial laboratory habitats and represent a restricted range of life histories, metabolic diversity and habitat types. Thus, they typically lack predictive power when applied to real ecosystems (e.g., Reusch and Wood 2007). Fortunately, recent technological and analytical advances in NGS are helping bridge the gap between extensive ecological knowledge about key species and the genetic base of variation in their phenotypic traits (e.g., Ekblom and Galindo 2011). Development in this direction will ultimately lead to an integrative view of genetics, physiology, performance, and fitness, as recommended to marine zooplanktologists by Arnold (1983) more than 30 years ago. Therefore, in this dissertation we initiate the development of *Calanus finmarchicus*, an extensively studied, ecologically important copepod, as an integrative model for ecological and evolutionary studies.

Most of the ecologically important traits of *Calanus*, such as fitness, survival, reproduction, and distribution, are largely dependent on sea temperature (Kjellerup et al. 2012; Helaouët and Beaugrand 2007; Hirche 1987, 1997). While the real temperature optimum, where abundance peaks, is between 4°C and 7°C (Paper III; Sundby 2000; Wilson et al. 2015), the range -1.8 °C – 15 °C is seen as the realised niche (Helaouët et al. 2011), where growth, development and egg production increase with temperature (Campbell et al. 2001; Kjellerup et al. 2012; Preziosi and Runge 2014). However, at higher temperatures (up to 23 °C), fitness and reproduction drop, and survival is threatened (Hirche 1987). Interestingly, transcriptome-wide responses of *C. finmarchicus* to temperature range not only matches the optimum temperature of the species (Paper III), but also showed that minor deviations from optimal temperature for a long time can initiate changes in gene expression, before any physiological

changes would be detected (Paper III). The sensitivity of transcriptomic analysis can be utilized to investigate the influence of other environmental factors, including toxicants, natural (e.g., toxic diatom bloom; Lauritano et al. 2011)) and industrial pollution (e.g., oil and diesels; Hjorth and Nielsen 2011; Hansen et al. 2013a).

Precise identification of the closely related species, *C. finmarchicus* and *C. glacialis*, based on new morphological and molecular methods (Paper I and II), allowed species-level comparative study of transcriptomic responses. The observed contrasting responses between the two species (Paper III) is likely to influence the fundamental niches of the species and their species-specific adaptations, and raise new questions regarding molecular pathways or regulatory mechanisms that may be involved in thermal stress responses of the Arctic *C. glacialis*.

Phenotypic plasticity together with genetic adaptation and underlying genetic variation influence the ability of individuals and populations to cope with changing and stressful environmental conditions, and thereby, determine fitness (Bijlsma and Loeschcke 2012). Investigation of transcriptome-wide responses to stress (Paper III) could be seen as the foundation for future research aimed at detection of adaptation and plasticity, and comparing expression of target genes responding to the same stressors in different populations over the distributional range of a species. A first such attempt was to compare transcriptomic thermal stress responses between Arctic (Disko Bay) (Paper III) and mid-Atlantic populations of *C. finmarchicus* (Smolina et al. unpublished), which showed that more genes were responding to stress in the Arctic population. Similar population specificity in the expression of stress-associated genes between southern and northern populations was observed in copepod *Tigriopus californicus* (Schoville et al. 2012), killifish *Fundulus heteroclitus* (Whitehead and Crawford 2006), and sea grass *Zostera marina* (Bergmann et al. 2010; Franssen et al. 2011). As in many other cases (e.g., Bergmann et al. 2010; Schoville et al. 2012), at this stage is its difficult to separate local adaptation and phenotypic plasticity in *C. finmarchicus*, since evidences exist that favour either mechanism. For instance, evidence of genetic differentiation between populations of *C. finmarchicus* has been detected both within and between basins of the North Atlantic (Paper III; Unal and

Bucklin 2010), which could be a result of selection or lead to locally adapted populations. In addition, locally adapted populations have been documented for *Tigriopus californicus* (Schoville et al. 2012; Willet 2010), *Acartia hudsonica* (Colin and Dam 2002), and *Calanus helgolandicus* (Lauritano et al. 2012). Nevertheless, evidence from a 50-year time series in the North Sea suggests lack of thermal adaptation in *C. finmarchicus,* leading to dramatic distributional changes (Helaouët and Beaugrand 2007; Hinder et al. 2014).

On the other hand, phenological plasticity may play a substantial role in thermal stress responses. For instance, increase in acclimation temperature could expand the thermal tolerance of several copepods, including the congeneric *Calanus sinicus* (Jiang et al. 2008). This could explain the weaker gene expression response in the mid-Atlantic population of *C. finmarchicus* with *in situ* temperature of 10 ˚C compared to the Arctic population with *in situ* temperature of 0 ˚C (Paper III; Smolina et al. unpublished). Moreover, Halcrow (1963) showed that *C. finmarchicus* do not acclimate to temperatures outside their seasonal temperature range, but can compensate for temperature changes close to the *in situ* water temperatures at the collection location. However more detailed studies, including experimental ones, similar to those performed for *Daphnia magna*, a freshwater crustacean with a wide temperature range of habitat (Yampolsky et al. 2014), are needed to investigate contributions of both genetic adaptation and phenotypic plasticity to thermal tolerance of *C. finmarchicus.* In addition, both phenotypic plasticity and genetic adaptation will be important in adaptive responses to ongoing climate change (e.g., Munday et al. 2013), and the potential for both must be estimated to understand evolutionary potential of *C. finmarchicus* and other *Calanus* species.

Temperature is also suggested to control duration of another key life history trait of copepods in the genus *Calanus*: diapause (Pierson et al. 2013). And while both, control of the onset of diapause and its duration, is an active area of research, particularly on molecular level (Aruda et al. 2011; Tarrant et al. 2008, 2014), investigation of different populations and under different temperature conditions is missing. Linking diapause to genetic variation and adaptation could clarify whether diapause is altered by natural

selection, as has been shown for some zooplankton (Avery 2005; Marcus 1984), and how it will change with changes in climate.

The effects of temperature, as a single factor, upon organisms could be modified by simultaneous action of several environmental factors (e.g., salinity, acidification, hypoxia). The interaction of factors could have neutral, synergistic or antagonistic characteristics (e.g., Duman et al. 2014; Madeira et al. 2014; Torstensson et al. 2013). Thus, for better understanding of population persistence, more studies are needed to investigate influences of natural environmental variability (e.g., Unal et al. 2013) and mixtures of environmental stressors (e.g., Hansen et al. 2007). Studying these environmental interactions will present logistical and conceptual challenges to the study of evolutionary adaptation of zooplankton to global change (Dam 2013). In addition, the impact of a factors may vary among developmental stages, as was detected in *C. finmarchicus* for temperature (Kvile et al. 2014), carbon dioxide concentration (Cripps et al. 2014), and the breadth of the ecological niche (Reygondeau and Beaugrand 2010). Thus, investigated transcriptomic responses to thermal stress (Paper III) is female-biased and additional stage-specific investigations are needed to understand population dynamics and persistence.

Knowledge about the whole set of optimum environmental conditions for different *Calanus* species could be obtained by investigations of their distributional ranges and vertical migration. However, traditional morphological methods of *Calanus* identification can be biased (e.g., Gabrielsen et al. 2012; Parent et al. 2011, 2012), due to co-occurrence of *Calanus* species (Conover 1988; Kwasniewski et al. 2003; Wilson et al. 2015) and their similar morphology and overlapping sizes (Frost 1974; Kwasniewski et al. 2003; Wilson et al. 2015). A new InDel panel of genetic markers could solve the problem in a simple, reliable, and high-throughput way for many individuals of any development stage (Paper I). The approach could also provide an alternative to the usual pooling of early developmental stages into one *Calanus* spp. category, due to the impossibility of correct species identification (eg., Blachowiak-Samolyk et al. 2006; Hirche and Kwasniewski 1997), and thus, yield important ecological information about larval and juvenile stages. In addition, many monitoring cruises collect zooplankton

samples in formalin, which makes samples unsuitable for some types of genetic identification methods. Since DNA extraction from formalin samples results in short DNA fragments, amplification of long mitochondrial (Bucklin et al. 1995b; Hill et al. 2001; Lindeque et al. 1999) or microsatellite (Parent et al. 2012; Provan et al. 2007) amplicons will be impeded. On the other hand, the InDel markers have rather short amplicon sizes and have already yielded positive results with DNA from formalin samples (Smolina et al. unpublished), while these markers lack drawbacks of other existing markers (reviewed in 1.2.1). The new genetic markers open a door for investigation of samples from current and historical cruises. Overall, the association of particular species with specific environmental conditions will allow further investigation of the effects of environmental change on zooplankton species with different physiological and biological properties (Lindeque et al. 1999).

3.3. *Calanus* as a model for evolutionary genetics

Species of the copepod genus *Calanus* show potential as model species for studies of evolutionary genetics of marine copepods, including speciation, hybridization and interspecific genetic differentiation. Speciation in *Calanus* has likely occurred with a recent radiation (possibly during the late Pleistocene) that kept morphology similar (Fleminger and Hulsemann 1977). Moreover, 14 *Calanus* species (Bucklin et al. 1995b) appear to be adapted to different marine environments with large differences in temperature preference – occupying waters from equator to both poles (Conover 1988; Jaschnov 1970; Kosobokova and Hirche 2001; Pakhomov and McQuaid 1996; Uye 2000). Therefore, adaptation to environmental habitat, particularly to temperature, may be integral to the mechanisms of speciation (Keller and Seehausen 2012), which are likely to be associated with palaeoceanographic events and resulted in geographic isolation, genetic divergence, and reproductive isolation (Frost 1974; Fleminger and Hulsemann 1977). *Calanus* shows large variation in genome size (from ≈ 6.34 to ≈ 12.19 Gb) (McLaren et al. 1988), considerable genetic sequence divergence in the mitochondrial region of 16S rRNA (Bucklin et al. 1995a) and across genomes (e.g., Paper I), and species-specific metabolites (Hansen et al. 2013b). Investigation of the

transcriptomes of two closely related species, *C. finmarchicus* and *C. glacialis* revealed that they are relatively conserved between these species (Paper I, III), while regulation of gene expression in response to stress can be very different (Paper III; Hansen et al. 2011, 2013). Since altered gene regulation is often involved into rapid evolution (Doebley and Lukens 1998; Ferea et al. 1999; Van Laere et al. 2003), it can be hypothesized that this may play a role in evolution between closely related *Calanus* species. It would be interesting to test this hypothesis by analysing transcriptome-wide responses to stress in *Calanus* species. Sister species pairs that are adapted to different environments could be of particular interest, although pairs living in similar conditions, but in isolation, e.g., in polar (*C. hyperboreus* and *C. propinquus*) and temperate regions (*C. finmarchicus* and *C. sinicus* or *C. pacificus*), may serve as interesting examples of possible parallel adaptation. Some clues about parallel evolution in *Calanus* could also be obtained by investigating the patterns of mutation rate variation and selection intensity (d_N / d_S ratio) at orthologous genes across genomes of *Calanus* species, as was performed for 10 avian species (Kunstner et al. 2010).

Lack of complete reproductive isolation between species could lead to hybridization, which is common in the marine environment (Gardner 1997; Roques et al. 2001; Uthicke et al. 2005), but less so in marine planktonic taxa. Nevertheless, hybrids between *C. finmarchicus* and *C. glacialis* may serve as an example (Parent et al. 2012). Hybrids of the two *Calanus* species were not significantly different from those of their maternal ancestor for prosome length, red pigmentation level, and the egg production rate (Parent et al. 2015). Although the hybrids appear to be fertile and were abundant at some stations in Canadian Arctic (Parent et al. 2012), both parental species still dominated the hybrid zone (Parent et al. 2012, 2015). Thus, the fitness of hybrids is probably lower than that of the parental species over their entire life cycle, and further, the effect of hybridization on species population dynamics is rather minor (Parent et al. 2015). Nevertheless, discrimination of pure species and hybrids is essential for ecological or evolutionary research. Surprisingly, among individuals used for this dissertation and several side projects (in total over 1300 individuals of *C. finmarchicus* and *C. glacialis*) no hybrids were detected using nuclear InDel markers

(Paper I), while a few could be classified as hybrids using microsatellites loci (according to Paper II). One possible explanation could be that hybrids are artefacts of the cross-species application of *C. finmarchicus* microsatellites (e.g., Curtu et al. 2004). This is supported by the fact that most of the hybrid traits were similar to their maternal species, including metabolism and the egg membrane (Parent et al. 2015). It is also possible that our samples contained only pure species, however sampling for this dissertation was performed across the distributional range of *C. finmarchicus,* including areas of possible hybridization with *C. glacialis,* where the time difference between the reproductive periods of the species is short (e.g., Disko Bay, western Greenland; Madsen et al. 2001). Nevertheless, experimental crosses and investigation of larger sub-sections of the genome using NGS technologies will allow address the question via genome-wide hybridization and introgression studies in *Calanus* species, despite limited genome sequence data (e.g., Twyford and Ennos 2012). In addition to possible hybridization between *C. finmarchicus* and *C. glacialis*, possible hybridization between *C. glacialis* and *C. marshallae* should be investigated, since these sister species have very similar sequences for some marker genes (Bucklin et al. 1995a,b; Paper I) and have more similar genome size (McLaren et al. 1988).

At the within-species level, most of studies suggest genetic differentiation between populations of *C. finmarchicus* both within and between ocean basins (Paper IV; reviewed in section 1.2.3.). These findings are supported by significant variation in life history traits, such as numbers of generations per year, timing of reproduction and diapause, and seasonal patterns of abundance (Heath et al. 2000, 2004; Hirche 1996; Melle et al. 2014a,b; Planque et al 1997; Tande 1991). The evolutionary mechanisms driving such differentiation remain largely unexplored, but two non-exclusive explanations are emerging: (1) entrainment into currents/gyres or (2) different selection pressures acting in the different regions. In the marine environment, ocean currents, while providing a means for connectivity, also produce eddies, fronts and gyres that act as retention mechanisms for pelagic eggs, larvae and zooplankton (Banks et al. 2007; Knutsen et al. 2007; Mackas et al. 2005). In addition, modelling studies have shown that ocean current velocity and seasonal changes in current

structure can affect dispersal in meroplankton between sites (White et al. 2010), and two adjacent sites may rarely exchange migrants if they are located on different sides of an oceanographic front (Gilg and Hilbish 2003). Thus, ocean currents may act as efficient barriers to gene flow, similar to continental landmasses (e.g., Peijnenburg et al. 2004; Goetze 2005, 2011; Papadopoulos et al. 2005; Yebra et al. 2011). Several copepod species display genetically distinct populations between gyres within ocean basins, indicating that physical retention may be an important process promoting genetic structure (Andrews et al. 2014; Goetze 2005, 2011; Norton and Goetze 2013). In the North Atlantic, modelling studies have shown a high degree of connectivity between the gyres over a time-scale of six years, resulting in transport of *C. finmarchicus* populations between the two main gyre systems (i.e., Labrador/Irminger Sea and Norwegian Sea; Speirs et al. 2006). However, this model assumed that transport is purely passive (Speirs et al. 2006), thus overlooking active behaviours such as diel and ontogenetic vertical migration of *C. finmarchicus* (Dale and Kaartvedt 2000; Unstad and Tande 1991) that may result in retention of copepods within a gyre system.

On the other hand, even if connectivity is sufficient to transport *C. finmarchicus* throughout the whole North Atlantic, different selection pressures could lead to isolation by adaptation (Nosil et al. 2009). For instance, the western, central and eastern North Atlantic are characterised by different seasonally and geographically varying ranges of temperature, salinity and light conditions, which provide a variety of habitats for its biota (Melle et al. 2014b). Thus, habitat-specific selection may be a mechanism driving differentiation, particularly as selection is proposed to be widespread in marine zooplankton and likely to be a dominant driver of evolution (Peijnenburg and Goetze 2013). For example, adaptation to a specific salinity and/or temperature regime has been suggested to be a driver of genetic diversification in copepods (Chen and Hare 2008, 2011; Yebra et al. 2011), shrimps (Jorde et al. 2015), and chaetognaths (Peijnenburg et al. 2004). Although our tests for selection using > 300 SNPs were negative among six populations of *C. finmarchicus* (Paper IV), there is evidence suggesting that selection may drive differentiation of locally adapted

northern and southern populations (section 3.2.). Moreover, if selection is an important driver of genetic differentiation, its strength and the effects will differ between unlinked genetic markers, leading to large differences in the degree of population structure detected by different loci (Peijnenburg and Goetze 2013; Piganeau et al. 2011). This could account for the discrepancy between differentiation based on microsatellite and SNP markers in *C. finmarchicus* (Paper IV; Provan et al. 2008; Unal and Bucklin 2010). To some extent, similar cases of discrepancy in the observed strength of differentiation have been reported for the planktonic chaetognath *Sagitta setosa* (Peijnenburg et al. 2006), copepod *Haloptilus longicornis* (Andrews et al. 2014; Norton and Goetze 2013), and the Patagonian toothfish *Dissostichus eleginoides* (Shaw et al. 2004), although based on comparisons between nuclear microsatellite markers and mitochondrial DNA. Other evidence that suggests selection could underlay large differences between census and effective population sizes (Peijnenburg and Goetze 2013). The effective population size in *C. finmarchicus* was suggested to be 10^7 to 10^8 times smaller than the census size (Bucklin and Wiebe 1998). Application of genome-wide SNP-based NGS approaches (genome scan) could help identify genomic regions that are likely to be under selective pressure, and therefore address questions of adaptive differentiation between populations (Lexer et al. 2013). This may help to recognise the action of environmental agent(s), as well as predict impacts of climate change (Stillman and Armstrong 2015) using, for instance, time-series samples from Continuous Plankton Recorder surveys.

3.4. Implications for management

Calanus finmarchicus is considered to be a potentially huge biological resource, with annual biomass production of about 300 million tonnes in the Nordic Seas (Skjoldal 2004), but currently there is no commercial harvest of this resource. The wax ester-rich oil of *C. finmarchicus* can be used as an effective supplement to fish diets (e.g., Olsen et al. 2004), leading to better growth and nutrient utilization efficiency (Colombo-Hixson et al. 2013). In the near future, the oil extracted from *C. finmarchicus* may also be utilized in the health care industry as health- and nutrition products.

Indeed, preclinical tests using rodent models have demonstrated that *C. finmarchicus* oil has the ability to reduce body weight, inflammation, atherosclerotic plaque formation and visceral fat deposition (Höper et al. 2011; Pedersen et al. 2014). In addition, the effects of *Calanus* oil were shown to be not only preventive, but also therapeutic (Höper et al. 2013). The prospect of large-scale commercial harvesting is thus growing (Tacon and Metian 2015) and caution should be applied, since *C. finmarchicus* is crucial for the ecosystems, including important fish stocks (Beaugrand and Reid 2003; Beaugrand et al. 2003; Gislason and Astthorsson 2002; Søreide et al. 2008;). Thus, if *C. finmarchicus* in fact comprises several semi-isolated populations (Paper IV; Unal and Bucklin 2010), fisheries management regulations should be applied or commercial culturing should be considered. Moreover, in light of climate change, it might be necessary to consider climate-driving shifts of *C. finmarchicus* distribution northward (Beaugrand et al. 2002; Chust et al. 2013), and therefore reduce pressure on more southerly populations, which are likely to be more vulnerable (Hinder et al. 2014).

4. Conclusion

With ecosystems facing more and more anthropogenic and climate-related challenges, in depth understanding of keystone species is crucial to predict ecosystem responses. The present research fills several knowledge gaps and contributes to a better understanding of the ecology and genetics of *Calanus* species and their potential to cope with climate change. *Calanus* species identification, thermal stress response, and population genetic structure, have been long-standing subjects of scientific research and in the present dissertation I have addressed these questions using state-of-the-art next generation sequencing technologies. The main outcomes are: the development of robust genetic markers for *Calanus* species identification, identification and comparison of transcriptome–wide responses to thermal stress in two closely related *C. finmarchicus* and *C. glacialis*, and detection of genetically differentiated populations of *C. finmarchicus*. Altogether, these findings lead to better understanding of ecological and genetic features of *Calanus* species, and consequently, of its past (e.g., speciation and evolution), present (e.g., adaptation and hybridization), and future (e.g., response to climate change).

5. Future perspectives – need for a genome

The research presented in this dissertation uses advances of NGS to obtain and compare homologous genomic and transcriptomic sequences (Paper I), to observe dynamic changes of expression levels of transcripts (Paper III), and to detect SNP variation at many genomic loci (Paper IV). These applications were possible due to partial *de novo* assembly of obtained sequencing reads. However, the data analysis was challenging and resulted in reduced amounts of information from the sequencing, due to the unavailability of at least draft of *Calanus* genome. Indeed, a reference genome for RNA-seq would allow detection of more transcripts, better discrimination between expression of paralogous genes, and result in clearer answers regarding genomic changes that are occurring during population responses to stressful environment (Stillman and Armstrong 2015). For population genetic structure, a reference genome would be a great advantage as well. For instance, a reference genome can result in precise prediction of ddRAD fragments (e.g., Lepais and Weir 2014) leading to a greater number of loci that could be explored for variation, a primary drawback for Paper IV. Moreover, knowledge of the genome could be used to target specific genomic regions via Cap-seq (section 1.3.4), which has also potential to succeed with historical samples collected in formalin. On the other hand, comparing genomic variation between populations of *C. finmarchicus* exhibiting different numbers of generations per year, timing of reproduction and diapause over its geographic distribution (Hirche 1996a,b; Melle et al. 2014; Planque et al 1997; Tande 1991) may also result in detection of candidate genes that regulate life history traits (Fabian et al. 2012).

A *Calanus* reference genome could greatly assist in inferring gene homology within a species (paralogy: gene duplication leading to multiple gene copies) and across species (orthology: genes descended from the same ancestral sequence), which is difficult when dealing with *de novo* assemblies of genome or transcriptome. Both *C. finmarchicus* and *C. glacialis* appear to have diversified gene families (Paper III; Lenz et al. 2014). Among them, families of different *hsp* genes seem to be one of the most complex and important for survival under stress conditions (Hofmann and Todgham

2010). For example, an increase in copy numbers of *hsp* genes and differential regulation may play an important role in thermoresistance of Diptera species (Garbuz et al. 2011). Overall, gene duplication and differential expression of paralogs can be important for flexible phenotypic responses to ecological challenges (Colbourne et al. 2011). Therefore, identification of the whole *hsp* repertoire in *Calanus* species and assessment of its regulation under native and stressful conditions will be important for understanding intra- and interspecies tolerance and adaptation.

The contrasting thermal stress responses between *C. finmarchicus* and *C. glacialis* shown by differences in gene expression (Paper III) lead to questions about putative mechanisms of stress response regulation. Epigenetics, for example, can also influence variation in the stress response (Boyko and Kovalchuk 2008; Crews et al. 2012). Epigenetic variation reflects changes at the molecular level, which do not change underlying DNA sequence, but regulate gene expression via DNA methylation, histone modifications, chromatin remodeling, and expression of small non-coding RNAs such as microRNA (miRNA) (Bossdorf et al. 2008). Thus, epigenetic mechanisms can affect ecologically important traits even in the absence of genetic variation, and play a crucial role for an organism's immediate and evolutionary response to its environment (Kilvitis et al. 2014). For instance, genetically identical lines of *Arabidopsis thaliana* that differed in DNA methylation showed heritable phenotypic variation and plasticity in ecologically relevant traits (Zhang et al. 2013). Such epigenetic variation could be considered as non-genetic "plasticity loci" (Schlichting and Pigliucci 1993), and play a particularly important role in rapid adaptation to changing environments. Moreover, epigenetic marks can be passed on to the subsequent generations, and may be able to produce genetic change (e.g., through the regulation of transposable elements or an effect on recombination), thus, contributing to long-term evolution via adaptation and ecological speciation (reviewed in Smith and Ritchie 2013). Long restricted to model species, investigation of epigenetic mechanisms is starting on ecologically important species. This endeavour has led to a new field, ecological epigenetics, which attempts to understand the role and significance of epigenetic processes in the context of ecology and evolution (Bossdorf et al. 2008; Richards et al. 2012; Smith and Ritchie

2013). *Calanus* could be an interesting system for investigation of ecological epigenetics, due to its broad environmental habitats, variation in life history traits and rapid, possibly ecological, speciation; however, a reference genome is an imperative requirement for such studies.

6. References

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 Biol Ecol 461:479-488.

Albrechtsen A, Nielsen FC, Nielsen R (2010) Ascertainment biases in SNP chips affect measures of population divergence. Mol Biol Evol 27:2534–2547.

Anderson EC, Skaug HJ, Barshis DJ (2014) Next-generation sequencing for molecular ecology: a caveat regarding pooled samples. Mol Ecol 23:502-512.

Andrews KR, Norton EL, Fernandez-Silva I, Portner E, Goetze E (2014) Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. Mol Ecol 23:5462-5479.

Angeloni F, Wagemaker N, Vergeer P, Ouborg J (2012) Genomic toolboxes for conservation biologists. Evol Appl 5:130-143.

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K (2013) RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. Mol Ecol 22:3179-3190.

Arnold SJ (1983) Morphology, performance and fitness. Am Zool 23:347–361.

Aruda AM, Baumgartner MF, Reitzel AM, Tarrant AM (2011) Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*. J Insect Physiol 57:665-675.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al. (2000) Gene Ontology: tool for the unification of biology. Nature Genetics, 25, 25–29.

Avery DE (2005) Induction of embryonic dormancy in the calanoid copepod *Acartia hudsonica*: heritability and phenotypic plasticity in two geographically separated populations. J Exp Mar Biol Ecol 314:215–225.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3:e3376.

Baker M (2012) *De novo* genome assembly: what every biologist should know. Nat Methods 9:333-337.

Banks SC, Piggott MP, Williamson JE, Bové U, Holbrook NJ, Beheregaray LB (2007) Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. Ecol 88:3055–3064.

Beaugrand G (2004) The North Sea regime shift: evidences, causes, mechanisms and consequences. Prog Oceanography 60:245 – 262.

Beaugrand G (2009) Decadal changes in climate and ecosystems in the North Atlantic Ocean and adjacent seas. Deep-Sea Res II 56: 656–673.

Beaugrand G (2014) Pelagic ecosystems and climate change. In: Freedman B (Eds), Global environmental change. Springer Netherlands, pp. 141-150.

Beaugrand G, Brander KM, Lindley JA, Souissi S, Reid PC (2003) Plankton effect on cod recruitment in the North Sea. Nature 426:661-664.

Beaugrand G, Conversi A, Chiba S, Edwards M, Fonda-Umani S, Greene C et al. (2015) Synchronous marine pelagic regime shifts in the Northern Hemisphere. Phil. Trans. R. Soc. B 370: 20130272.

Beaugrand G, Edwards M, Legendre L (2010) Marine biodiversity, ecosystem functioning, and carbon cycles. PNAS 107:10120-10124.

Beaugrand G, Mackas D, Goberville E (2013) Applying the concept of the ecological niche and a macroecological approach to understand how climate influences zooplankton: Advantages, assumptions, limitations and requirements. Prog Oceanography 111:75-90.

Beaugrand G, Reid F, Ibañez F, Lindley A, Edwards M (2002) Reorganization of North Atlantic marine copepod biodiversity and climate. Sci 296:1692–1694.

Beaugrand G, Reid PC (2003) Long-term changes in phytoplankton, zooplankton and salmon related to climate. Glob Chang Biol 9:801-817.

Berge J, Gabrielsen TM, Moline M, Renaud PE (2012) Evolution of the Arctic *Calanus* complex: an Arctic marine avocado? J Plankton Res 34:191-195.

Bergmann N, Winters G, Rauch G, Eizaguirre C, Gu J, Nelle P, et al. (2010) Population-specificity of heat stress gene induction in northern and southern eelgrass *Zostera marina* populations under simulated global warming. Mol Ecol 19:2870-2883.

Bijlsma R, Loeschcke V (2012) Genetic erosion impedes adaptive responses to stressful environments. Evol Appl 5:117-129.

Blachowiak-Samolyk K, Kwasniewski S, Richardson K, Dmoch K, Hansen E, Hop H et al. (2006) Arctic zooplankton do not perform diel vertical migration (DVM) during periods of midnight sun. Mar Ecol Prog Ser 308:101-116.

Bonnet D, Richardson A, Harris R, Hirst A (2005) An overview of *Calanus helgolandicus* ecology in European waters. Prog Oceanography 65:1-53.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. Ecol Lett 11:106-115.

Boyko A, Kovalchuk I (2008) Epigenetic control of plant stress response. Environ Mol Mutagenesis 49:61–72.

Bron JE, Frisch D, Goetze E, Johnson SC, Lee CE, Wyngaard GA (2011) Observing copepods through a genomic lens. Front Zool 8:22.

Bucklin A, Astthorsson OS, Gislason A, Allen LD, Smolenack SB, Wiebe PH (2000a) Population genetic variation of *Calanus finmarchicus* in Icelandic waters: preliminary evidence of genetic differences between Atlantic and Arctic populations. ICES J Mar Science 57:1592–1604.

Bucklin A, Frost B, Kocher T (1995a) DNA sequence variation of the mitochondrial 16S rRNA in *Calanus* (Copepoda; Calanoida): intraspecific and interspecific patterns. Mol Mar Biol Biotechnol 1:397-407.

Bucklin A, Frost B, Kocher T (1995b) Molecular systematics of six *Calanus* and three *Metridia* species (Calanoida, Copepoda). Mar Biol 121:655–664.

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

Bucklin A, Kocher TD (1996) Source regions for recruitment of *Calanus finmarchicus* to Georges Bank: evidence from molecular population genetic analysis of mtDNA. Deep-Sea Res 43:1665–1682.

Bucklin A, Lajeunesse TC (1994) Molecular genetic variation of *Calanus pacificus* (Copepoda: Calanoida): preliminary evaluation of genetic structure and subspecific differentiation based on mtDNA sequences. Rep CCOFI 35:45–51.

Bucklin A, Sundt R, Dahle G (1996) Population genetics of *Calanus finmarchicus* in the North Atlantic. Ophelia 44:29–45.

Bucklin A, Wiebe PH (1998) Low mitochondrial diversity and small effective population sizes of the copepods *Calanus finmarchicus* and *Nannocalanus minor*: Possible impact of climatic variation during recent glaciation. J Heredity 89:383-392.

Campana MG, Robles García NM, Tuross N (2015) America's red gold: multiple lineages of cultivated cochineal in Mexico. Ecol Evol 5:607-617.

Campbell RG, Wagner MW, Teegarden GJ, Boudreau CA, Durbin EG (2001) Growth and development rates of the copepod *Calanus finmarchicus* in the laboratory. Mar Ecol Prog Ser 221:161–183.

Carstensen J, Weidmann A, Olszewska A, Kwasniewski S (2012) Effects of environmental conditions on the biomass of *Calanus* spp. in the Nordic Seas. J Plankton Res 34:951–966.

Chan YC, Roos C, Inoue-Murayama M, Inoue E, Shih CC, Pei KJC, et al. (2010) Mitochondrial genome sequences effectively reveal the phylogeny of *Hylobates gibbons*. PLoS One 5:e14419.

Chen G, Hare MP (2008) Cryptic ecological diversification of a planktonic estuarine copepod, *Acartia tonsa*. Mol Ecol 17:1451–1468.

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.

Choi J-W, Liao X, Park S, Jeon H-J, Chung W-H, Stothard P, et al. (2013) Massively parallel sequencing of Chikso (Korean brindle cattle) to discover genome-wide SNPs and InDels. Moll Cells 36:203-211.

Chu ND, Kaluziak CT, Trussell GC, Vollmer SV (2014) Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus*. Mol Ecol 23:1863-1873.

Chust G, Castellani C, Licandro P, Ibaibarriaga L, Sagarminaga Y, Irigoien X (2014) Are *Calanus* spp. shifting poleward in the North Atlantic? A habitat modelling approach. ICES J Mar Sci 71:241–253.

Clément JAJ, Toulza E, Gautier M, Parrinello H, Roquis D, Boissier J et al. (2013) Private selective sweeps identified from next generation pool-sequencing reveal convergent pathways under selection in two inbred *Schistosoma mansoni* strains. PLoS Negl Trop Dis 7:e2591.

Colbourne JK, Pfrender ME, Gilbert D, Thomas WK, Tucker A, Oakley TH et al: The ecoresponsive genome of *Daphnia pulex*. Sci 331:555-561.

Colin SP, Dam HG (2002) Latitudinal differentiation in the effects of the toxic dinoflagellate *Alexandrium* spp. on the feeding and reproduction of populations of the copepod *Acartia hudsonica*. Harmful Algae 1:113–125.

Colombo-Hixson SM, Olsen RE, Tibbetts SM, Lall SP (2013) Evaluation of *Calanus finmarchicus* copepod meal in practical diets for juvenile *Atlantic halibut* (Hippoglossus hippoglossus). Aquaculture Nutr 19:687-700.

Conover RJ (1988) Comparative life histories in the genera *Calanus* and *Neocalanus* in high latitudes of the northern hemisphere. Hydrobiologia 167 (168):127–142.

Cook N, Aziz N, Hedley PE, Morris J, Milne L., Karley AJ et al. (2011) Transcriptome sequencing of an ecologically important graminivorous sawfly: a resource for marker development. Conservation Genet Resour 3:789–795.

Crews D, Gillette R, Scarpino SV, Manikkam M, Savenkov MI, Skinner MK (2012) Epigenetic transgenerational inheritance of altered stress responses. PNAS 109:9143-9148.

Cripps G, Lindeque P, Flynn KJ (2014) Have we been underestimating the effects of ocean acidification in zooplankton? Glob Chang Biol 20:3377-3385.

Curtu A, Finkeldey R, Gailing O (2004) Comparative sequencing of a microsatellite locus reveals size homoplasy within and between European oak species (*Quercus* spp.). Plant Mol Biol Reporter 22:339–346.

Dale T, Kaartvedt S (2000) Diel patterns in stage-specific vertical migration of *Calanus finmarchicus* in habitats with midnight sun. ICES J Mar Sci 57:1800–1818.

Dalpadado P, Ingvaldsen RB, Stige LC, Bogstad B, Knutsen T, Ottersen G et al. (2012) Climate effects on Barents Sea ecosystem dynamics. ICES J Mar Sci 69:1303-1316.

Dam HG (2013) Evolutionary adaptation of marine zooplankton to global change. Annu Rev Mar Sci 5:349-370.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499-510.

Dawson MN, Hamner WM (2005) Rapid evolutionary radiation of marine zooplankton in peripheral environments. PNAS 102:9235–9240.

De Wit P, Palumbi SR (2013) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol 22:2884-2897.

De Wit P, Pespeni MH, Ladner JT, Barshis DJ, Seneca F, Jaris H et al. (2012) The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. Mol Ecol Resour 12:1058-1067.

Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell 10:1075–1082.

Doney SC, Ruckelshaus M, Duffy JM, Barry JP, Chan F, English CA et al. (2012) Climate change impacts on marine ecosystems. Annu Rev Mar Sci 4:11–37.

Drinkwater KF, Beaugrand G, Kaeriyama M, Kim S, Ottersen G, Perry RI, et al. (2010) On the processes linking climate to ecosystem changes. J Mar Syst 79:374-388.

Duman F, Koca FD, Sahan S (2014) Antagonist effects of sodium chloride on the biological responses of an aquatic plant (*Ceratophyllum demersum* L.) exposed to hexavalent chromium. Water Air Soil Pollut 225:1865.

Edwards M, Beaugrand G, Helaouët P, Alheit J, Coombs S (2013) Marine ecosystem response to the Atlantic Multidecadal Oscillation. PLoS One 8:e57212.

Edwards M, Richardson AJ (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. Nature 430:881–884.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107:1-15.

Ellertsen B, Melle W (2011) Secondary production in the Norwegian Sea. Havforskningsrapporten 2011 (in Norwegian). Fisken og Havet, særnummer 1-2011:84–85.

Fabian DK, Kapun M, Nolte V, Kofler R, Schmidt PS, Schlötterer C et al. (2012) Genome-wide patterns of latitudinal differentiation among populations of *Drosophila melanogaster* from North America. Mol Ecol 21:4748-4769.

Faircloth BC (2008) msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8: 92–94.

Falk-Petersen S, Mayzaud P, Kattner G, Sargent JR (2009) Lipids and life strategy of Arctic *Calanus*. Mar Biol Resour 5:18-39.

Falk-Petersen S, Pavlov V, Timofeev S, Sargent JR (2007) Climate variability and possible effects on arctic food chains: the role of *Calanus*. In: Ørbæk JB, Kallenborn R, et al. (Eds), Arctic alpine ecosystems and people in a changing environment. Springer-Verlag, Berlin, pp. 147–166.

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. Annu Rev Physiol 61:243-282.

Ferea TL, Botstein D, Brown PO, Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. PNAS 96:9721–9726.

Ferretti L, Ramos-Onsins SE, Pérez-Enciso M (2013) Population genomics from pool sequencing. Mol Ecol 22:5561-5576.

Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. PNAS 105:17994–17999.

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

Franssen SU, Gu J, Bergmann N, Winters G, Klostermeier UC, Rosenstiel P, et al. (2011) Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a marine foundation species. PNAS 108:19276-19281.

Frost B (1974) *Calanus marshallae*, a new species of Calanoid copepod closely allied to the sibling species *C. finmarchicus* and *C. glacialis*. Mar Biol 26: 77–99.

Funari V, Canosa SJ (2014) The Importance of bioinformatics in NGS: Breaking the bottleneck in data interpretation. Sci 344:653.

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

Gabrielsen T, Merkel B, Søreide J, Johansson-Karlsson E, Bailey A, Vogedes D et al. (2012) Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. Polar Biol 35:1621–1628.

Gao Q, Jin K, Ying S-H, Zhang Y, Xiao G, Shang Y et al. (2011) Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet 7:e1001264.

Garbuz D, Yushenova I, Zatsepina O, Przhiboro AA, Bettencourt BR, Evgenév M (2011) Organization and evolution of *hsp70* clusters strikingly differ in two species of stratiomyidae (Diptera) inhabiting thermally contrasting environments. BMC Evol Biol 11:74.

Gardner JPA (1997) Hybridization in the sea. In: Blaxter HS, Southward AJ (Eds), Advances in Marine Biology. Academic Press Limited New York, pp. 1-78.

Gautier M, Foucaud J, Gharbi K, Cézard T, Galan M, Loiseau A et al. (2013a) Estimation of population allele frequencies from next-generation sequencing data: pool-versus individual-based genotyping. Mol Ecol 22:3766-3779.

Gautier M, Gharbi K, Cezard T, Foucaud J, Kerdelhué C, Pudlo P (2013b) The effect of RAD allele dropout on the estimation of genetic variation within and between populations. Mol Ecol 22:3165-3178.

Gifford DJ, Harris RP, McKinnell SM, Peterson WT, St John MA (2010) Dynamics of marine ecosystems: target species. In: Barange M, Harris R, Field J, Hoffman E, Perry I, Werner C. (Eds.), Global change and marine ecosystems. Oxford University Press, pp. 75–88.

Gilg MR, Hilbish TJ (2003) The geography of marine larval dispersal: coupling genetics with fine-scale physical oceanography. Ecol 84:2989–2998.

Gilg O, Kovacs KM, Aars J, Fort J, Gauthier G, Grémillet D et al. (2012) Climate change and the ecology and evolution of Arctic vertebrates. Ann NY Acad Sci 1249:166–190.

Gislason A, Astthorsson OS (2002) The food of Norwegian spring-spawning herring in the western Norwegian Sea in relation to the annual cycle of zooplankton. Sarsia 87:236–247.

Goberville E, Beaugrand G, Edwards M (2014) Synchronous response of marine plankton ecosystems to climate in the Northeast Atlantic and the North Sea. J Mar Syst 129:189-202.

Godden GT, Jordon-Thaden IE, Chamala S, Crowl AA, García N, Germain-Aubrey CC et al. (2012) Making next-generation sequencing work for you: approaches and practical considerations for marker development and phylogenetics. Plant Ecol Diversity 5:427-450.

Goetze E (2003) Cryptic speciation on the high seas; global phylogenetics of the copepod family *Eucalanidae*. Proc R Soc Lond B 270:2321–2331.

Goetze E (2005) Global population genetic structure and biogeography of the oceanic copepods *Eucalanus hyalinus* and *E. spinifer*. Evol 59:2378-2398.

Goetze E (2011) Population differentiation in the Open Sea: Insights from the pelagic copepod *Pleuromamma xiphias*. Integr Comp Biol 51:580–597.

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformat 29:1072-1075.

Halcrow K (1963) Acclimation to temperature in the marine copepod, *Calanus finmarchicus* (Gunner.). Limnol Oceanogr 8:1-8.

Hansen BH, Altin D, Nordtug T, Olsen AJ (2007) Suppression subtractive hybridization library prepared from the copepod *Calanus finmarchicus* exposed to a sublethal mixture of environmental stressors. Comp Biochem Physiol D 2:250-256.

Hansen BH, Altin D, Øverjordet IB, Jager T, Nordtug T (2013a) Acute exposure of water soluble fractions of marine diesel on Arctic *Calanus glacialis* and boreal *Calanus finmarchicus*: Effects on survival and biomarker response. Sci Total Environment 449:276-284.

Hansen BH, Degnes K, Øverjordet IB, Altin D, Størseth TR (2013b) Metabolic fingerprinting of arctic copepods *Calanus finmarchicus*, *Calanus glacialis* and *Calanus hyperboreus*. Polar Biol 36:1577-1586.

Hansen J, Sato M, Ruedy R, Lo K, Lea DW, Medina-Elizade M (2006) Global temperature change. PNAS 103:14288-14293.

Harris V, Edwards M, Olhede SO (2014) Multidecadal Atlantic climate variability and its impact on marine pelagic communities. J Mar Syst 133:55-69.

Harvey MG, Smith BT, Glenn TC, Faircloth BC, Brumfield RT (2013) Sequence capture versus restriction site associated DNA sequencing for phylogeography. arXiv:1312.6439.

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish 9:333–362.

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). Prog Oceanogr 59:1–30.

Head EJH, Melle W, Pepin P, Bagøien E, Broms C (2013) On the ecology of *Calanus finmarchicus* in the Subarctic North Atlantic: A comparison of population dynamics and environmental conditions in areas of the Labrador Sea-Labrador/Newfoundland Shelf and Norwegian Sea Atlantic and Coastal Waters. Prog Oceanogr 114:46–63.

Heath MR, Astthorsson OS, Dunn J, Ellertsen B, Gaard E, Gislason A, Gurney WSC, et al. (2000) Comparative analysis of *Calanus finmarchicus* demography at locations around the Northeast Atlantic. ICES J Mar Sci 57:1562-1580.

Heath MR, Backhaus JO, Richardsson K, McKenzie E, Slagstad D, Beare D, et al. (1999) Climate fluctuations and the spring invasion of the North Sea by *Calanus finmarchicus*. Fish Oceanogr 8:163–176.

Heath MR, Boyle PR, Gislason A, Gurney WSC, Hay SJ, Head EJH, et al. (2004) Comparative ecology of over-wintering *Calanus finmarchicus* in the northern North Atlantic, and implications for life-cycle patterns. ICES J Mar Sci 61:698-708.

Heath MR, Rasmussen J, Ahmed Y, Allen J, Anderson CIH, Brierley AS, et al. (2008) Spatial demography of *Calanus finmarchicus* in the Irminger Sea. Prog Oceanogr 76:39–88.

Helaouët P, Beaugrand G (2007) Macroecology of *Calanus finmarchicus* and *C. helgolandicus* in the North Atlantic Ocean and adjacent seas. Mar Ecol Prog Ser 345:147–165.

Helaouët P, Beaugrand G (2009) Physiology, ecological niches and species distribution. Ecosystems 12:1235–1267.

Helaouët P, Beaugrand G, Reid PC (2011) Macrophysiology of *Calanus finmarchicus* in the North Atlantic Ocean. Prog Oceanogr 91:217–228.

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT et al. (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Resour 11:123-136.

Hickling R, Roy DB, Hill JK, Fox R, Thomas CD (2006) The distributions of a wide range of taxonomic groups are expanding polewards. Glob Chang Biol 12:450– 455.

Hill R, Allen L, Bucklin A (2001) Multiplexed species-specific PCR protocol to discriminate four N. Atlantic *Calanus species*, with an mtCOI gene tree for ten *Calanus* species. Mar Biol 139:279–287.

Hinder SL, Gravenor MB, Edwards M, Ostle C, Bodger OG, Lee PLM et al. (2014) Multi-decadal range changes vs. thermal adaptation for north east Atlantic oceanic copepods in the face of climate change. Glob Chang Biol 20:140-146.

Hirche H-J (1987) Temperature and plankton II. Effect on respiration and swimming activity in copepods from the Greenland Sea. Mar Biol 94:347-356.

Hirche H-J (1997) Life cycle of the copepod *Calanus hyperboreus* in the Greenland Sea. Mar Biol 128:607-617.

Hirche H-J, Kosobokova K (2007) Distribution of *Calanus finmarchicus* in the northern North Atlantic and Arctic Ocean—Expatriation and potential colonization. Deep-sea Res II 54:2729-2747.

Hirche H, Hagen W, Mumm N, Richter C (1994) The northeast water Polynya, Greenland Sea. III. Mesozooplankton and macrozooplankton distribution and production of dominant herbivorous copepods during spring. Polar Biol 14:491–503.

Hirche HJ, Brey T, Niehoff B (2001) A high frequency time series at Ocean Weather Ship Station M (Norwegian Sea): population dynamics of *Calanus finmarchicus*. Mar Ecol Prog Ser 219:205–219.

Hirche HJ, Kwasniewski S (1997) Distribution, reproduction and development of *Calanus* species in the Northeast water in relation to environmental conditions. J Mar Syst 10:299-317.

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

Hjorth M, Nielsen TG (2011) Oil exposure in a warmer Arctic: potential impacts on key zooplankton species. Mar Biol 158:1339-1347.

Hoegh-Guldberg O, Bruno JF (2010) The impact of climate change on the world's Marine Ecosystems. Sci 328:1523-1528.

Hofmann GE, Todgham AE (2010) Living in the now: physiological mechanisms to tolerate a rapidly changing environment. Annu Rev Physiol 72:127-145.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6: e1000862.

Höper AC, Salma W, Khalid AM, Hafstad AD, Sollie SJ, et al. (2013) Oil from the marine zooplankton *Calanus finmarchicus* improves the cardiometabolic phenotype of diet-induced obese mice. Br J Nutr 110:2186-93.

Höper AC, Salma W, Sollie SJ, Hafstad AD, Lund J, Khalid AM, et al. (2014) Wax esters from the marine copepod *Calanus finmarchicus* reduce diet-induced obesity and obesity-related metabolic disorders in mice. J Nutr 144:164-169.

Huang Y, Liu G, Chen X (2014) Molecular phylogeography and population genetic structure of the planktonic copepod *Calanus sinicus* Brodsky in the coastal waters of China. Acta Oceanol Sin 33:74-84.

Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS et al. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods 12:115-121.

Hunter ME, Hart KM (2013) Rapid microsatellite marker development using next generation pyrosequencing to inform invasive Burmese python—*Python molurus bivittatus*—Management. Int J Mol Sci 14:4793-4804.

Imelfort M, Duran C, Batley J, Edwards D (2009) Discovering genetic polymorphisms in next-generation sequencing data. Plant Biotechnol J 7:312–317.

Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC et al. (2008) Genotype, haplotype and copy-number variation in worldwide human populations. Nat 451:998.

Jaschnov WA (1970) Distribution of *Calanus* species in the seas of the Northern hemisphere. Internationale Revue der gesamten Hydrobiologie und Hydrographie 55:197–212.

Ji R, Yu H, Fu Q, Chen H, Ye W, Li S, et al. (2013) Comparative transcriptome analysis of salivary glands of two populations of rice brown planthopper, *Nilaparvata lugens*, that differ in virulence. PLoS One 8:e79612.

Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, et al. (2011) Synthetic spike-in standards for RNA-seq experiments. Genome Res 21:1543–1551.

Jiang Z, Zeng J, Chen Q, Huang Y, Xu X, Liao Y, et al. (2008) Tolerance of copepods to short-term thermal stress caused by coastal power stations. J Therm Biol 33:419-423.

Jorde PE, Søvik G, Westgaard J-I, Albretsen J, André C, Hvingel C et al. (2015) Genetically distinct populations of northern shrimp, *Pandalus borealis*, in the North Atlantic: adaptation to different temperatures as an isolation factor. Mol Ecol 24:1742-1757.

Jump AS, Peñuelas J (2005) Running to stand still: adaptation and the response of plants to rapid climate change. Ecol Lett 8:1010-1020.

Kann LM, Wishner K (1996) Genetic population structure of the copepod *Calanus finmarchicus* in the Gulf of Maine: allozyme and amplified mitochondrial DNA variation. Mar Biol 125:1432–1793.

Karam M-J, Lefèvre F, Dagher-Kharrat MB, Pinosio S, Vendramin GG (2014) Genomic exploration and molecular marker development in a large and complex conifer genome using RADseq and mRNAseq. Mol Ecol Res 15:601-612.

Karlsen BO, Klingan K, Emblem Å, Jørgensen TE, Jueterbock A, Furmanek T (2013) Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. Mol Ecol 22:5098-5111.

Keller I, Seehausen O (2012) Thermal adaptation and ecological speciation. Mol Ecol 21:782–799.

Kendziorski C, Irizarry RA, Chen KS, Haag JD and others (2005) On the utility of pooling biological samples in microarray experiments. PNAS 102:4252-4257

Kenny EM, Cormican P, Gilks WP, Gates AS, O'Dushlaine CT, Pinto C et al. (2011) Multiplex target enrichment using DNA indexing for ultra-high throughput SNP detection. DNA Res 18:31–38.

Kilvitis HJ, Alvarez M, Foust CM, Schrey AW, Robertson M, Richards CL (2014) Ecological epigenetics. Ecol Genomics. Adv Exp Med Biol 781:191-210.

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.

Kjellerup S, Dünweber M, Swalethorp R, Nielsen TG, Møller EF, Markager S et al. (2012) The effect of a future warmer ocean on the co-existing copepods *Calanus finmarchicus* and *C. glacialis* in the Disko bay western Greenland. Mar Ecol Prog Ser 447:87-108.

Knowlton N (1993) Sibling species in the sea. Ann Rev Ecol Syst 24:189–216.

Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420:73–90.

Knutsen H, Olsen EM, Ciannelli L, Espeland SH, Knutsen AJ, Simonsen JH, et al. (2007) Egg distribution, bottom topography and small-scale cod population structure in a coastal marine system. Mar Ecol Prog Ser 333:249–255.

Konczal M, Koteja P, Stuglik MT, Radwan J, Babik W (2014) Accuracy of allele frequency estimation using pooled RNA-Seq. Mol Ecol Resour 14:381-392.

Kosobokova KN, Hirche HJ (2001) Reproduction of *Calanus glacialis* in the Laptev Sea, Arctic Ocean. Polar Biol 24:33–43.

Kozol R, Blanco-Bercial L, Bucklin A (2012) Multi-gene analysis reveals a lack of genetic divergence between *Calanus agulhensis* and *C. sinicus* (Copepoda; Calanoida). PLoS One 7:e45710.

Kristiansson E, Asker N, Förlin L, Larsson DJ (2009) Characterization of the *Zoarces viviparus* liver transcriptome using massively parallel pyrosequencing. BMC Genomics 10:345.

Künstner A, Wolf JBW, Backström N, Whitney O, Balakrihnan CN, Day L et al. (2010) Comparative genomics based on massive parallel transcriptome sequencing reveals patterns of substitution and selection across 10 bird species. Mol Ecol 19:266-276.

Kvile KØ, Dalpadado P, Orlova E, Stenseth NC, Stige LC (2015) Temperature effects on *Calanus finmarchicus* vary in space, time and between developmental stages. Mar Ecol Prog Ser 517:85-104.

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

Lauritano C, Borra M, Carotenuto Y, Biffali E, Miralto A, Procaccini G, et al. (2011) Molecular evidence of the toxic effects of diatom diets on gene expression patterns in copepods. PLoS ONE 6:e26850.

Lauritano C, Carotenuto Y, Miralto A, Procaccini G, Ianora A (2012) Copepod population-specific response to a toxic diatom diet. PLoS ONE 7:e47262.

Lenz PH, Roncalli V, Hassett RP, Wu L-S, Cieslak MC, Hartline DK, et al. (2014) *De novo* assembly of a transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) – the dominant zooplankter of the North Atlantic Ocean. PLoS ONE 9:e88589.

Lenz PH, Unal E, Hassett RP, Smith CM, Bucklin A, Christie AE et al (2012) Functional genomics resources for the North Atlantic copepod, *Calanus finmarchicus*: EST database and physiological microarray. Comp Biochem Physiol D 7:110-123.

Lepais O, Weir JT (2014) SimRAD: an R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. Mol Ecol Resour 14:1314-1321.

Lexer C, Mangili S, Bossolini E, Forest F, Stolting KN, Pearman PB, et al. (2013) "Next generation" biogeography: Towards understanding the drivers of species diversification and persistence. J Biogeograph 40:1013–1022.

Lindeque P, Harris R, Jones M, Smerdon GR (1999) Simple molecular method to distinguish the identity of *Calanus* species (Copepoda: Calanoida) at any developmental stage. Mar Biol 133, 91–96.

Lindeque P, Hay S, Heath M, Ingvarsdottir A, Rasmussen J, Smerson GR et al. (2006) Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. J Plankton Res 28:221–238.

Liu S, Gao G, Palti Y, Cleveland BM, Weber GM, Rexroad CE (2014) RNA-seq analysis of early hepatic response to handling and confinement stress in rainbow trout. PLoS One 9:e88492.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

Mackas DL, Tsurumi M, Galbraith MD, Yelland DR (2005) Zooplankton distribution and dynamics in a North Pacific Eddy of coastal origin: II. Mechanisms of eddy colonization by and retention of offshore species. Deep Sea Res II: Topical Studies in Oceanography 52:1011–1035.

Madeira D, Narciso L, Diniz MC, Vinagre C (2014) Synergy of environmental variables alters the thermal window and heat shock response: An experimental test with the crab *Pachygrapsus marmoratus*. Mar Environ Res 98:21–28.

Madsen SD, Nielsen TG, Hansen BW (2001) Annual population development and production by *Calanus finmarchicus*, *C. glacialis* and *C. hyperboreus* in Disko Bay, western Greenland. Mar Biol 139:75–93.

Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A et al. (2009) Target-enrichment strategies for next- generation sequencing. Nat Methods 7:111–118.

Marcus NH (1984) Variations in the diapause response of *Labidocera aestiva* (Copepoda: Calanoida) from different latitudes and its importance to the evolutionary process. Biol Bull 166:127–139.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. Mol Phylogenet Evol 66:526-538.

McLaren IA, Sevigny J-M, Corkett CJ (1988) Body sizes, development rates, and genome sizes among *Calanus* species. Biology of Copepods 47:275-284.

Melle W, Runge J, Head E, Plourde S, Castellani C, Licandro P, et al. (2014a) Biogeography of key mesozooplankton species in the North Atlantic, by manual counting methods, and egg production of *Calanus finmarchicus*. Earth Syst Sci Data Discuss 7:225–242.

Melle W, Runge J, Head E, Plourde S, Castellani C, Licandro P, et al. (2014b) The North Atlantic Ocean as habitat for *Calanus finmarchicus*: Environmental factors and life history traits. Prog Oceanogr 129:244-284.

Meng T (2003) Studies on the feeding of anchovy (*Engraulis japonicus*) at different life stages on zooplankton in the Middle and Southern Waters of the Yellow Sea. Mar Fish Res (in Chinese) 24:1–9.

Metzker ML (2010) Sequencing technologies — the next generation. Nat Rev Genet 11:31-46.

Meyer M, Stenzel U, Hofreiter M (2008) Parallel tagged sequencing on the 454 platform. Nat Protoc 3:267–278.

Michaud J, Taggart CT (2007) Lipid and gross energy content of North Atlantic right whale food, *Calanus finmarchicus*, in the Bay of Fundy. Endang Species Res 3:77–94.

Miner BE, De Meester L, Pfrender ME, Lampert W, Hairston NG (2012) Linking genes to communities and ecosystems: *Daphnia* as an ecogenomic model. Proc R Soc B 279:1873-1882.

Munday PL, Warner RR, Monro K, Pandolfi JM, Marshall DJ (2013) Predicting evolutionary responses to climate change in the sea. Ecol Lett 16:1488-1500.

Narum SR, Buerkle SA, DAVEY JW, Miller MR, Hohenlohe PA (2013) Genotyping-by-sequencing in ecological and conservation genomics. Mol Ecol 22:2841-2847.

Nelson RJ, Carmack EC, McLaughlin FA, Cooper GA (2009) Penetration of Pacific zooplankton into the western Arctic Ocean tracked with molecular population genetics. Mar Ecol Prog Ser 381:129-138.

Niehoff B, Hirche HJ (1996) Oogenesis and gonad maturation in the copepod *Calanus finmarchicus* and the prediction of egg production from preserved samples. Polar Biol 16: 601–612.

Niehoff B, Madsen S, Hansen BW, Nielsen TG (2002) Reproductive cycles of three dominant *Calanus* species in Disko Bay, West Greenland. Mar Biol 140:567–576.

Ning J, Wang M, Li C, Sun S (2013) Transcriptome sequencing and *de novo* analysis of the copepod *Calanus sinicus* using 454 GS FLX. PLoS ONE 8:e63741.

Nolte V, Pandey RV, Kofler R, Schlötterer C (2013) Genome-wide patterns of natural variation reveal strong selective sweeps and ongoing genomic conflict in *Drosophila mauritiana*. Genome Res 23:99-110.

Norris RD (2000) Pelagic species diversity, biogeography, and evolution. Paleobiol 26:236–58.

Norton EL, Goetze E (2013) Equatorial dispersal barriers and limited population connectivity among oceans in a planktonic copepod. Limnol Oceanogr 58:1581-1596.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18:375–402.

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A et al. (2008) Genes mirror geography within Europe. Nat 456:98.

Nuwer ML, Frost BW, Armbrust EV (2008) Population structure of the planktonic copepod *Calanus pacificus* in the North Pacific Ocean. Mar Biol 156:107-115.

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto Encyclopedia of genes and genomes. Nucleic Acids Research, 27, 29–34.

Olsen EM, Ottersen G, Llope M, Chan K, Beaugrand G, Stenseth NC (2011) Spawning stock and recruitment in the North Sea cod shaped by food and climate. Proc R Soc Lond 278:504–510.

Olsen RE, Henderson RJ, Sountama J, Hemre GI, Ringø E, Melle W, et al. (2004) Atlantic salmon, *Salmo salar*, utilizes wax ester-rich oil from *Calanus finmarchicus* effectively. Aquaculture 240:433-449.

Pakhomov EA, McQuaid CD (1996) Distribution of surface zooplankton and seabirds across the Southern Ocean. Polar Biol 16:271-286.

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Annu Rev Ecol Syst 25:547-572.

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.

Parent G, Plourde S, Turgeon J (2011) Overlapping size ranges of *Calanus* spp. off the Canadian Arctic and Atlantic Coasts: impact on species' abundances. J Plankton Res 33, 1654–1665.

Parent G, Plourde S, Turgeon J (2012) Natural hybridization between *Calanus finmarchicus* and *C. glacialis* (Copepoda) in the Arctic and North-west Atlantic. Limnol Oceanogr 57:1057–1066.

Parent GJ, Plourde S, Joly P, Turgeon J (2015) Phenology and fitness of *Calanus glacialis*, *C. finmarchicus* (Copepoda), and their hybrids in the St. Lawrence Estuary. Mar Ecol Prog Ser 524:1-9.

Parmesan C (2006) Ecological and evolutionary response to recent climate change. Ann Rev Ecol Evol Syst 37:637–669.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37–42.

Pasternak A, Riser CW, Arashkevich E, Rat'kova T, Wassmann P (2002) *Calanus* spp. grazing affects egg production and vertical carbon flux (the marginal ice zone and open Barents Sea). J Mar Sys 38:147–164.

Pedersen AM, Salma W, Höper AC, Larsen TC, Olsen RL (2014) Lipid profile of mice fed a high-fat diet supplemented with a wax ester-rich marine oil. Eur J Lipid Sci Technol 116:1718-1726.

Peijnenburg K, Breeuwer JAJ, Pierrot-Bults AC, Menken SBJ (2004) Phylogeography of the planktonic chaetognath *Sagitta setosa* reveals isolation in European seas. Evol 58:1472–1487.

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.

Peijnenburg KTCA, Goetze E (2013) High evolutionary potential of marine zooplankton. Ecol Evol 3:2765-2781.

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 147:1270-1289.

Pérez-Porro AR, Navarro-Gómez D, Uriz MJ, Giribet G (2013) A NGS approach to the encrusting Mediterranean sponge *Crella elegans* (Porifera, Demospongiae, Poecilosclerida): transcriptome sequencing, characterization and overview of the gene expression along three life cycle stages. Mol Ecol Resour 13:494-509.

Perry AL, Low PJ, Ellis JR, Reynolds JD (2005) Climate change and distribution shifts in marine fishes. Sci 308:1912–1915.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135.

Pevzner PA, Tang HX, Waterman MS (2001) An Eulerian path approach to DNA fragment assembly. PNAS 98:9748–9753.

Pierson JJ, Batchelder H, Saumweber W, Leising A, Runge J (2013) The impact of increasing temperatures on dormancy duration in *Calanus finmarchicus*. J Plankton Res 35:504-512.

Piganeau G, Eyre-Walker A, Grimsley N, Moreau H (2011) How and why DNA barcodes underestimate the diversity of microbial eukaryotes. PLoS ONE 6:1–6.

Planque B, Hays GC, Ibanez F, Gamble JC (1997) Large scale spatial variations in the seasonal abundance of *Calanus finmarchicus*. Deep-Sea Res 44:315–326.

Pörtner HO, Farrell AP (2008) Physiology and climate change. Sci 322:690–692.

Preziosi BM, Runge JA (2014) The effect of warm temperatures on hatching success of the marine planktonic copepod, *Calanus finmarchicus*. J Plankton Res 36:1381-1384.

Provan J, Beatty G, Keating S, Maggs C, Savidge G (2008) High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus*. Proc R Soc Lond Ser B 276:301–307.

Provan J, Beatty G, Maggs C, Savidge G (2007) Expressed sequence tag-derived microsatellites for the cool-water marine copepod *Calanus finmarchicus*. Mol Ecol Notes 7:1369–1371.

Ramos AA, Weydmann A, Cox CJ, Canário AVM, Serrão EA, Pearson GA (2015) A transcriptome resource for the copepod *Calanus glacialis* across a range of culture temperatures. Mar Genomics doi:10.1016/j.margen.2015.03.014.

Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P et al. (2013) Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol 14:R95.

Reid PC, 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. Fish Oceanogr 12:260–269.

Reusch TBH (2013) Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. Evol Appl 7:104-122.

Reusch TBH, Wood TE (2007) Molecular ecology of global change. Mol Ecol 16:3973-3992.

Reygondeau G, Beaugrand G (2010) Water column stability and *Calanus finmarchicus*. J P Res 33:119–136.

Reygondeau G, Beaugrand G (2011) Future climate-driven shifts in distribution of *Calanus finmarchicus*. Glob Chang Biol 17:756-766.

Richards CL, Verhoeven KJF, Bossdorf O (2012) Evolutionary significance of epigenetic variation. In: Wendel JF (Eds), Molecular biology and evolution of the plant genome. Springer Verlag, Vienna/New York, pp 257–274.

Richardson AJ (2008) In hot water: zooplankton and climate change. ICES J Mar Sci 65:279–295.

Richardson AJ, Field JG, Fowler JL, Mitchell-Innes BA, Verheye HM (2003) Seasonal and event-scale variation in growth of *Calanus agulhensis* (Copepoda) in the Benguela upwelling system and implications for spawning of sardine *Sardinops sagax*. Mar Ecol Prog Ser 254:239–251.

Riesgo A, Andrade SCS, Sharma PP, Novo M, Perez-Porro AR, Vahtera V, et al. (2012) Comparative description of ten transcriptomes of newly sequenced invertebrates and efficiency estimation of genomic sampling in non-model taxa. Front Zool 9: 33.

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformat 26:139-140.

Rokas A, Abbot P (2009) Harnessing genomics for evolutionary insights. Trends Ecol Evol 24:192-200.

Roques S, Sevigny JM, Bernatchez L (2001) Evidence for broadscale introgressive hybridization between two redfish (genus *Sebastes*) in the North-west Atlantic: a rare marine example. Mol Ecol 10:149-165.

Rossetto M, Henry RJ (2014) Escape from the laboratory: new horizons for plant genetics. Trends Plant Sci 19:554-555.

Ruperao P, Edwards D (2014) Bioinformatics: Identification of markers from next-generation sequence data. In: Batley J (Eds), Plant Genotyping. Methods and Protocols 1245:29-47.

Santure AW, Stapley J, Ball AD, Birkhead TR, Burke T, Slate J (2010) On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. Mol Ecol 19:1439–1451.

Schizas NV, Dahms H-U, Ricaurte M, Hwang J-S (2014) Population genetic patterns of the copepod *Calanus sinicus* in the northwest Pacific. Challenges Aquat Sci 741:153-165.

Schlichting CD, Pigliucci M (1993) Control of phenotypic plasticity via regulatory genes. Am Naturalist 142:366−370.

Schliesky S, Gowik U, Weber APM, Bräutigam A (2012) RNA-seq assembly – are we there yet? Front Plant Sci 3:1-12.

Schlötterer C (2004) The evolution of molecular markers - just a matter of fashion? Nat Rev Genet 5:63-69.

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.

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS (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.

Shaw PW, Arkhipkin AI, Al-Khairulla H (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. Mol Ecol 13:3293–3303.

Simon SA, Zhai J, Nandety RS, McCormick KP, Zeng J, Mejia D et al. (2009) Short-read sequencing technologies for transcriptional analyses. Ann Rev Plant Biol 60:305–333.

Singhal S (2013) De novo transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. Mol Ecol Resour 13:403-416.

Skjoldal HR, Dalpadado P, Dommasnes A (2004) Food webs and trophic interactions. In: Skjoldal HR (Eds), The Norwegain sea ecosystem. Tapir Academic Press, Trondheim, pp. 447–506.

Slagstad D, Ellingsen IH, Wassmann P (2011) Evaluating primary and secondary production in an Arctic Ocean void of summer sea ice: an experimental simulation approach. Prog Oceanogr 90:117–131.

Slate J, Gratten J, Beraldi D, Stapley J, Hale M, Pemberton J (2009) Gene mapping in the wild with SNPs: guidelines and future directions. Genetica 136:97–107.

Smith G, Ritchie MG (2013) How might epigenetics contribute to ecological speciation? Curr Zool 59:686–696.

Somero GN (2012) The physiology of global change: linking patterns to mechanisms. Annu Rev Mar Sci 4:39-61.

Soneson C, Delorenzi M (2013) A comparison of methods for differential expression analysis of RNA-seq data. BMC Bioinformat 14:91.

Søreide J, Falk-Petersen S, Nøst HE, Hop H, Carroll ML, Hobson (2008) Seasonal feeding strategies of *Calanus* in the high-Arctic Svalbard region. Deep Sea Res II 55:2225–2244.

Sørensen JG, Kristensen TN, Loeschcke V (2003) The evolutionary and ecological role of heat shock proteins. Ecol Lett 6:1025 –1037.

Southward AJ, Hawkins SJ, Burrows MT (1995) Seventy years' observations of changes in distribution and abundance of zooplankton and intertidal organisms in the western English Channel in relation to rising sea temperature. J Thermal Biol 20:127-155.

Speirs DC, Gurney WSC, Heath MR, Horbelt W, Wood SN, de Cuevas BA (2006) Ocean-scale modelling of the distribution, abundance, and seasonal dynamics of the copepod *Calanus finmarchicus*. Mar Ecol Prog Ser 313:173-192.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R et al. (2010) Adaptation genomics: the next generation. Trends Ecol Evol 25:705-712.

Stillman JH, Armstrong E (2015) Genomics are transforming our understanding of responses to climate change. BioSci doi: 10.1093/biosci/biu219.

Sundby S (2000) Recruitment of Atlantic cod stocks in relation to temperature and advection of copepod populations. Sarsia 85:277-298.

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 Fiord region, Spitsbergen. Variability Evol 2/3:113-119.

t Hoen PAC, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RHAM, de Menezes RX, et al. (2008) Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucl Acids Res 36: 1–11.

Tacon AGJ, Metian M (2015) Feed matters: satisfying the feed demand of aquaculture. Rev Fish Sci Aquaculture 23:1-10.

Tande KS (1991) *Calanus* in north Norwegian fjords and in the Barents Sea. Pol Res 10:389–407.

Tande KS, Miller CB (2000) Population dynamics of *Calanus* in the North Atlantic: results from the Trans-Atlantic study of *Calanus finmarchicus*. ICES J Mar Sci 57:1527.

Tarrant A, Baumgartner MF, Hansen BH, Altin D, Nordtug T, Olsen AJ (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.

Tarrant AM, Baumgartner MF, Verslycke T, Johnson CL (2008) Differential gene expression in diapausing and active *Calanus finmarchicus* (Copepoda). Mar Ecol Prog Ser 355:193-207.

Thackeray SJ, Jones ID, Maberly SC (2008) Long-term change in the phenology of spring phytoplankton: species-specific responses to nutrient enrichment and climatic change. J Ecol 96:523-535.

Thurston MI, Field D (2005) Msatfinder: detection and characterisation of microsatellites. Distributed by the authors athttp://www.genomics.ceh.ac.uk/msatfinder/. CEH Oxford: Oxford.

Tobler R, Franssen SU, Kofler R, Orozco-terWengel P, Nolte V, Hermisson J et al. (2014) Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. Mol Biol Evol 31:364–375.

Tomanek L (2010) Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. J Exp Biol 213:971-979.

Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR et al. (2013) ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ 1:e203.

Torstensson A, Hedblom M, Andersson J, Andersson MX, Wulff A (2013) Synergism between elevated pCO2 and temperature on the Antarctic sea ice diatom *Nitzschia lecointei*. Biogeosciences 10:6391-6401.

Toth AL, Varala K, Newman TC, Miguez FE, Hutchison SK, Willoughby DA et al. (2007) Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. Sci 318: 441–444.

Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV (2010) Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. Nat Genet 42: 260–263.

Twyford AD, Ennos RA (2012) Next-generation hybridization and introgression. Heredity 108:179-189.

Unal E, Bucklin A (2010) Basin-scale population genetic structure of the planktonic copepod *Calanus finmarchicus* in the North Atlantic Ocean. Prog Oceanography 87:175-185.

Unal E, Bucklin A, Lenz PH, Towle DW (2013) Gene expression of the marine copepod *Calanus finmarchicus*: Responses to small-scale environmental variation in the Gulf of Maine (NW Atlantic Ocean). J Exp Mar Biol Ecol 446:76-85.

Unal E, Frost BW, Armbrust V, Kideys AE (2006) Phylogeography of *Calanus helgolandicus* and the Black Sea copepod *Calanus euxinus*, with notes on *Pseudocalanus elongatus* (Copepoda, Calanoida). Deep-Sea Res II 53:1961-1975.

Unstad KH, Tande KS (1991) Depth distribution of *Calanus finmarchicus* and *C. glacialis* in relation to environmental conditions in the Barents Sea. Polar Res 10:409-420.

Uthicke S, Purcell S, Blockmans B (2005) Natural hybridization does not dissolve species boundaries in commercially important sea cucumbers. Biol J Linnean Soc 85:261-270.

Uye S (2000) Why does *Calanus sinicus* prosper in the shelf ecosystem of the Northwest Pacific Ocean? ICES J Mar Sci 57:1850–1855.

Van Laere A-S, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, et al. (2003) A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. Nature 425:832– 836.

Varpe Ø, Fiksen Ø, Slotte A (2005) Meta-ecosystems and biological energy transport from ocean to coast: the ecological importance of herring migration. Oecologia 146:443-451.

Wang M, Sun S, Li C, Shen X (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.

Wang S, Meyer E, McKay JK, Matz MV (2012) 2b-RAD: a simple and flexible method for genome-wide genotyping. Nat Methods 9:808–810.

Wassmann P, Duarte C, Agusti S, Sejr MK (2011) Footprints of climate change in the Arctic marine ecosystem. Glob Chang Biol 17:1235-1249.

Weslawski JM, Koszteyn J, Kwasniewski S, Stempniewicz L, Malinga M (1999) Summer food resources of the little auk, *Alle alle* (L.) in the European Arctic seas. Pol Polar Res 20:387–403.

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. Proc R Soc B: doi: 10.1098/rspb.2009.2214.

Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. PNAS 103:5425-5430.

Whitfield J (2008) Does 'junk food' threaten marine predators in northern seas? Sci 322:1786–1787.

Willet CS (2010) Potential fitness trade-offs for thermal tolerance in the intertidal copepod *Tigriopus californicus*. Evolution 64:2521–2534.

Wilson RG, Speirs DC, Heath MR (2015) On the surprising lack of differences between two congeneric calanoid copepod species, *Calanus finmarchicus* and *C. helgolandicus*. Prog Oceanogr 134:413-431.

Wolf JBW (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. Mol Ecol Resours 13:559-572.

Wolf JBW, Bayer T, Haubold B, Schilhabel M, Rosenstiel P, Tautz D (2010) Nucleotide divergence versus gene expression differentiation: comparative transcriptome sequencing in natural isolates from the carrion crow and its hybrid zone with the hooded crow. Mol Ecol 19:162–175.

Yampolsky LY, Schaer TMM, Ebert D (2014) Adaptive phenotypic plasticity and local adaptation for temperature tolerance in freshwater zooplankton. Proc R Soc B 281:20132744.

Yang Q, Sun F, Yang Z, Li H (2014) Comprehensive transcriptome study to develop molecular resources of the copepod *Calanus sinicus* for their potential ecological applications. BioMed Res International 2014: 493825.

Yebra L, Bonnet D, Harris RP, Lindeque PK, Peijnenburg KTCA (2011) Barriers in the pelagic: population structuring of *Calanus helgolandicus* and *C. euxinus* in European waters. Mar Ecol Prog Ser 428:135-149.

Young MD, McCarthy DJ, Wakefield MJ, Smyth GK, Oshlack A, Robinson MD (2012) Differential expression for RNA sequencing (RNA-Seq) data: mapping, summarization, statistical analysis, and experimental design. In: Rodríguez-Ezpeleta N et al. (Eds), Bioinformatics for high throughput sequencing. New York, Springer, pp. 169-190.

Zhang W, Chen J, Yang Y, Tang Y, Shang J, Shen B (2011) A practical comparison of *de novo* genome assembly software tools for next-generation sequencing technologies. PLoS One 6:e17915.

Zhang YY, Fischer M, Colot V, Bossdorf O (2013) Epigenetic variation creates potential for evolution of plant phenotypic plasticity. New Phytologist 197:314–322.