EURO-BASIN
BASIN SCALE ANALYSIS, SYNTHESIS AND INTEGRATION

SEVENTH FRAMEWORK PROGRAMME THEME 7 Environment

Collaborative project (Large-scale Integrating Project)

Project no: 246 933

Project Acronym: EURO-BASIN

Project title: European Basin-scale Analysis, Synthesis and Integration

Deliverable 3.5 Report on molecular genomic data for validation of basin-scale scenario models

Contributors: Galice Hoarau

Submission date: Oct 2014
Organisation name of the lead contractor of this deliverable: UiN

Start date of project: 31.12.2010 Duration: 48 months
Project Coordinator: Michael St John, DTU Aqua

Project co-funded by the European Commission within the Seventh Framework Programme, Theme 6 Environment

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission)	
RE	Restricted to a group specified by the consortium (including the Commission)	
CO	Confidential, only for members of the consortium (including the Commission)	

Deliverable 3.5 Report on molecular genomic data for validation of basin-scale scenario models, is a contribution to

Task 3.4: Broad-scale assessment of population genetics

Responsible: UiN; Participants: PML
Start month 1, end month 48

Executive Summary:

For the present task, we have relied on the use of state-of-the-art Next Generation Sequencing technology to 1) develop new molecular tools for *Calanus* species identification and 2) to investigate genetic structure of *Calanus* sp.

We have successfully developed a new set of InDel markers allowing the rapid and easy molecular identification of *Calanus* species. We have used these markers to investigate species boundaries in the North Atlantic. Contrary to what has been suggested before, we found no evidence for hybridisation between *C. finmarchicus* and *C. glacialis.*

For *Calanus finmarchicus,* both microsatellites and SNPs showed that there is no population structure across the entire North Atlantic. Genetic diversity is also homogeneous and relatively high in all populations. For *Calanus glacialis*, microsatellites markers also show a lack of population structure in the North Atlantic. Genetic diversity is also homogeneous but lower in comparison to *C. finmarchicus*. For both species it is thus likely that North Atlantic populations are part of single large panmictic units.

Relevance to the project & potential policy impact:

The markers for species ID have been used in WP4.

The lower genetic diversity found in *C. glacialis*, in combination with its apparent lack of heat shock response (see WP4) and the prediction for species distribution modelling (WP3 D3.8) suggest that *C. glacialis* populations in the North Atlantic are more vulnerable to climate change than *C. finmarchicus*.

Report:

Copepods of the genus *Calanus* play an important role both as consumers of primary production and as prey for many ecologically and commercially important species (Lønne & Gulliksen 1989). *Calanus* comprise up to 70 % of the zooplankton biomass (Head et al. 2003), with three species dominating the subarctic North Atlantic and Arctic Oceans: *Calanus finmarchicus, C. glacialis* and *C. hyperboreus* (Søreide et al. 2008; Falk-Petersen et al. 2009). These species are generally closely associated with the water masses where they originate from (Hirche & Kosobokova 2007), and therefore can be regarded as biological indicators of the North Atlantic sub-Arctic (*Calanus finmarchicus*) and Arctic (*C. glacialis* and *C. hyperboreus*) biomes (Kwasniewski et al. 2003; Daase et al. 2007). As a response to ocean warming, a northward shift of the subarctic species *C. finmarchicus* has been detected with a parallel decline throughout most of the North Atlantic (Beaugrand et al. 2002; Chust et al. 2014; Maar et al. 2013). A further northward shift of *C. finmarchicus* is predicted into Arctic waters that are dominated by *C. glacialis* and *C. hyperboreus* (Beaugrand et al. 2013; Wassmann et al. 2011), with a subsequent change in the food-web dynamics and secondary production (Falk-Petersen et al. 2007). Thus, *Calanus* species are central to many ecological, environmental and climatological studies (Gabrielsen et al. 2012) and correct identification of *Calanus* species and populations is essential.

Using traditional molecular markers (microsatellites) as well as state-of-the-art Next Generation Sequencing, we have focused on two main aspects; with part 1 focusing on species delineation and part 2 focusing on population genetic structure of *Calanus sp*.

PART 1 Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification

Introduction

Morphological identification of *Calanus* species is hampered by their subtle morphological differences, especially between early developmental stages (Hirche *et al.* 1994; Lindeque *et al.* 1999; Lindeque *et al.* 2006), resulting in consistent misidentification (Lindeque *et al.* 2006; Parent *et al.* 2011; Gabrielsen *et al.* 2012). Molecular identification of *Calanus* species is mainly based on mitochondrial markers, 16S rDNA (Bucklin *et al.* 1995; Lindeque *et al.* 1999) or cytochrome oxidase I (Hill *et al.* 2001). However, the recent report of hybrids among *Calanus* species (Parent *et al.* 2012) highlights the need for diagnostic nuclear marker in order to efficiently identify hybrids from parental species. Ten nuclear microsatellite loci have been developed for *C. finmarchicus* for in-depth population genetic studies (Provan et al. 2007), and have revealed hybridization between *C. finmarchicus* and *C. glacialis* in the Canadian Arctic (Parent *et al.* 2012). However, high mutation rate, lack of a mutation model, allelic dropout and difficulties to score alleles (for review see Pompanon *et al.* 2005; Selkoe & Toonen 2006), as well as homoplasy when comparing two species (Chambers & MacAvoy 2000; Curtu *et al.* 2004) can limit their use for species identification. Conversely, nuclear insertion/deletion (InDel) polymorphisms are co-dominant, have a low mutation rate, arise mainly from a single mutation event and provide a reasonably conserved phylogenetic signal (Liu & Cordes 2004; Nagy *et al.* 2012). The usability of InDel markers is further enhanced by ease of genotyping, repeatability, and possible use with degraded DNA (e. g. formalin or museum samples) due to generally small amplicon size (Väli *et al.* 2008; Pepinski *et al.* 2013).

Next generation sequencing (NGS) has revolutionized speed and availability of data acquisition in biology. For non-model species, NGS has dramatically facilitated the development of molecular markers such as microsatellites (Nie *et al.* 2012; Hunter & Hart 2013), InDels (Choi *et al.* 2013) and single nucleotide polymorphisms (SNPs) (Everett *et al.* 2011; Nussberger *et al.* 2013). Therefore, in the present study we relied on NGS to generate both genomic and transcriptomic sequences to develop a robust panel of nuclear InDel markers for the ease and versatile identification of *Calanus* spp.

Material and methods

Genome Ion Torrent PGM sequencing

Samples of *Calanus finmarchicus* and *C. glacialis* for genome sequencing were collected from two areas with minimum sympatry: *C. finmarchicus* close to Svinøy island (Norwegian Sea) and *C. glacialis* – in Rijpfjorden (Svalbard) (Table S1). Samples were collected using a WP2 net from 200m to the surface and stored in 95 % non-denatured ethanol. Preliminary identification of species was conducted by eye using prosome length criteria (Kwasniewski *et al.* 2003) and individuals with a prosome length far below/above the delimitation threshold were selected for further DNA extraction.

DNA was extracted individually using the E.Z.N.A.® Insect DNA Kit (Omega Bio-Tek) according to manufacturer´s instruction. To insure correct species identification, we used 6 microsatellite loci previously developed for *Calanus finmarchicus*: EL696609, EL585922, EH666870 (Provan *et al.* 2007); FK868270, FK670364, FK867682 (Parent *et al.* 2012). PCR reactions were performed in a total volume of 5 µl and contained 3 ng of DNA, reverse and fluorescently labeled forward primers with a final concentration of 0.25 µM/µl

each and 1X AmpliTaq Gold® PCR Master Mix (LifeTechnologies). A Veriti 96-Well Fast Thermal Cycler (LifeTechnologies) was used for PCR reactions with following cycling parameters: an initial denaturation step at 95 °C (10 min) followed by 40 cycles of 95 °C (20 s), 54 °C (for EL585922 and FK670364) or 56 °C (remaining 4 loci) (20 s), 72 °C (20 s) and a final extension at 72 °C (10 min). Fragment analysis was performed for all 6 loci in a single multiplex run on a 3500xL Genetic Analyzer (LifeTechnologies) following the manufacturer´s protocol. Alleles were scored using GENEMAPPER 3.7 (LifeTechnologies) and checked by eye. Species identification was based on results of STRUCTURE (v. 2.3.4) (Pritchard *et al*. 2000).

A single individual of each species was selected and 100 ng of DNA was subsequently used for library preparation. The Ion Plus Fragment Library Kit (Life Technologies) was used for the construction of 200 bp libraries according to manufacturer´s protocol. One library was constructed per species. Emulsion PCR was done in the Ion One Touch System, using the Ion One Touch 200 Template Kit (LifeTechnologies) followed by enrichment and quality control following the manufacturer´s protocol. Sequencing of each library was performed in the Ion Personal Genome Machine® (PGM™) System (LifeTechnologies) using the Ion PGM 200 Sequencing Kit and one Ion 316 chip (LifeTechnologies) per library according to manufacturer´s protocols.

Transcriptome Ion Torrent PGM sequencing

Samples for transcriptome sequencing were obtained from Disko Bay (West Greenland) (Table S1). Live females were morphologically identified using prosome length criteria and pigmentation/redness (Nielsen *et al*. submitted) and preserved in RNAlater (Qiagen). RNA and DNA were extracted simultaneously from each individual using the E.Z.N.A. DNA/RNA Isolation Kit (Omega Bio-Tek) according to manufacturer´s instruction.

Microsatellite genotyping was performed as previously described to ensure correct species identification. An equal amount of total RNA was pooled from 30 individuals per species. Pooled RNAs were treated with Baseline-ZERO DNase (Epicentre) and subsequently cleaned using the RNA Clean & Concentrator kit (Zymo Research, USA). Isolation of mRNA from total RNA was performed using two rounds of Mag-Bind mRNA Enrichment (Omega Bio-Tek, USA) with subsequent cleaning and concentration using the RNA Clean & Concentrator kit (Zymo Research, USA). Libraries were constructed using 40 ng of Poly(A) RNA and the Ion Total RNA-Seq Kit for AB Library Builder System (LifeTechnologies). Emulsion PCR was done in the Ion One Touch 2 System, using the Ion PGM Template OT2 200 Kit (LifeTechnologies) followed by enrichment and quality control following the manufacturer´s protocol. Sequencing of each library was performed in the Ion PGM System (LifeTechnologies) using the Ion PGM 200 Sequencing Kit and one Ion 318 chip (LifeTechnologies) for each species.

Bioinformatics

Sequences were analyzed using GENEIOUS 6.1.6 (Biomatters, available from http://www.geneious.com). Reads of genomes and transcriptomes for each species were analyzed separately using the following strategy. Reads were quality trimmed with an error probability limit of 0.05 and *de novo* assembled with medium-low sensitivity settings into four separate assemblies. To identify homologous regions between the two species, the 1000 longest genomic contigs and the 700 longest transcriptomic contigs of *C. finmarchicus* were used for BLAST searches (Megablast, default settings) against custom databases of genomic and transcriptomic contigs of *C. glacialis* respectively. Aligned regions between contigs were then visually scanned for large insertions/deletions between the two species

using the criteria: InDel size > 7 nucleotides for genome, ≥ 3 for transcriptome, and suitable conserved flanking regions for primer development. Primers were then designed using PRIMER3 (Koressaar & Remm 2007; Untergrasser *et al.* 2012) incorporated into GENEIOUS 6.1.6 (Biomatters).

Marker optimization and multiplexing

InDel optimization and multiplexing have been performed with DNA of two *C. finmarchicus* and two *C. glacialis* individuals identified as specimens for transcriptome Ion Torrent PGM sequencing (Table S1). PCR amplification of the candidate markers were tested using a Veriti 96-Well Fast Thermal Cycler (LifeTechnologies) in a total volume of 5 µl with 3 ng of DNA, 0.25 µM of each primer and 1X AmpliTaq Gold PCR Master Mix (LifeTechnologies). Cycling parameters included an initial denaturation step at 95 °C (10 min) followed by 40 cycles of 95 °C for 10 s, 55 °C for 10 s, 72 °C for 25 s, and a final extension at 72 °C for 7 min. Amplifications were checked on 2.7% agarose gels, and markers failing to amplify, showing multiple bands or no size difference between the two species were discarded. For the rest of the markers, new forward PCR primers were obtained with fluorescent labeling (VIC, 6-FAM, PET or NED) and tested using a 3500XL Genetic Analyzer (LifeTechnologies).

InDels test

To test validity of the new marker panel, 12 samples of *C. finmarchicus* and *C. glacialis* were obtained from both the West Greenland Sea and East Greenland Sea (Table S1). *Calanus* individuals were sampled with a WP2 net and preserved in 75 % non-denatured ethanol. Molecular identification based on a mitochondrial 16S rDNA fragment was performed according to Lindeque *et al.* (1999; 2006) with minor modifications and using half of the copepod. Modifications consisted of GoTAQ DNA polymerase (Promega) usage for PCR with the following cycling parameters: initial denaturation step at 94 °C (5 min) followed by 40 cycles of 94 °C (1 min), 45 °C (2 min), 72 °C (1 min), and a final annealing phase at 45 °C (2 min) and an extension phase at 72 °C (5 min). From the remaining half of each copepod DNA was extracted using the E.Z.N.A. Insect DNA Kit (Omega Bio-Tek) according to manufacturer´s instruction. Microsatellite loci were genotyped as described above. InDels were genotyped in four multiplex PCRs as follow: (1) G_150, G_155, G_461 and G_701; (2) T_595, T_1338 and T_6474; (3) T_1301 and T_1966; (4) T_4700, T_3133 and T_461. PCR conditions were as described above.

InDel sequencing and annotation

To ensure that homologous regions were amplified in both *C. finmarchicus* and *C. glacialis*, two individuals of each species (Table S1) were sequenced for each marker. PCR products were cleaned with ExoSAP-IT (Affymetrix), and sequenced using a BigDye Terminator v3.1 Cycle Sequencing Kit (LifeTechnologies) on a 3500XL Genetic Analyzer (LifeTechnologies) following the manufacturer´s protocols. Sequences were analyzed and aligned in GENEIOUS 6.1.6 (Biomatters). BLASTN searches against nucleotide and EST databases at NCBI (http://blast.ncbi.nlm.nih.gov) were used to annotate the regions/genes containing the InDels. Hits with E-value > 1e-10 were ignored.

Cross-amplification with other Calanus spp.

The usability of the 12 InDel markers was tested for *Calanus* species found in the North Atlantic and in the Arctic Oceans. DNA was extracted from two specimens of each of the following species: *C. hyperboreus* (Disko Bay), *C. helgolandicus* (Bay of Biscay) and *C.*

marshallae (Alaska) (Table S1) using the E.Z.N.A. Insect DNA Kit (Omega Bio-Tek). To ensure correct species identification we sequenced 400 bp of the 16S rDNA site using the following primers: C_16s_Fd1 (5´-GCCGCGTTAGTGYTAAGGTAGCA-3´) and C_16s_Rd1 (5´-AGAAACCAATCTGACTTRCGTCGA-3´). Amplification reactions were performed in a total volume of 10 µl with 6 ng of DNA, 0.25 µM of each primer and 1X AmpliTaq Gold PCR Master Mix (LifeTechnologies). Cycling parameters included an initial denaturation step at 95 °C (10 min) followed by 40 cycles of 95 °C, 62 °C, 72 °C for 20 s each step, and a final extension at 72 °C for 3 min. Amplifications were checked on 1 % agarose gel, and cleaned with ExoSAP-IT (Affymetrix). Sequencing was as described above and sequences were compared against nucleotide NCBI databases (http://blast.ncbi.nlm.nih.gov) using BLASTN. InDel amplification and genotyping was performed as described above.

Following the results of cross-species amplification, we tested different combinations of markers and amplification conditions to provide an easy and inexpensive protocol for high-throughput *Calanus* species identification in a single PCR.

Large scale genotyping

To investigate species boundaries, 300 females of *C. glacialis* and *C. finmarchicus* from Disko Bay, western Greenland and 89 females of *C. glacialis* and *C. finmarchicus* from Young Sound Fjord, Northeast Greenland were genotyped using the InDel markers. DNA extraction and genotyping was done as described above.

Results

For genome and transcriptome sequencing only individuals with morphological identification confirmed by nuclear microsatellites (probability of belonging to the cluster of either species at 99 %) were used. The sequencing of genomic libraries resulted in 3,118,080 reads for *C. finmarchicus* and in 1,920,496 reads for *C. glacialis*. After quality trimming, for *C. finmarchicus* 2,754,339 reads (mean length = 188 nt) were assembled into 36,134 contigs (mean contig length = 835 nt), and for *C. glacialis* 1,042,632 reads (mean length = 83.2 nt) were assembled into 52,597 contigs (mean contig length = 185 nt). The sequencing of transcriptome libraries of *C. finmarchicus* and *C. glacialis* resulted in 4,894,166 and 3,412,784 reads respectively. Quality trimming reduced number of reads to 3,548,728 (mean length = 108 nt) and 2,843,610 (mean length = 112 nt) respectively, and following assembly produced 251,042 (mean contig length = 225 nt) and 242,602 (mean contig length = 298 nt) contigs.

Primers were developed for 48 markers (InDel size ≥ 7nt) located in 43 genomic contigs and 31 markers (InDel size ≥ 3 nt) located in 30 transcriptomic contigs. Among those, four genome-based and eight transcriptome-based markers showed clear bands of a consistently different size between *C. finmarchicus* and *C. glacialis* (Table 1) and was consistent with other identification methods (Table S1). Sequencing the fragments for both species confirmed the homology of the amplicons. Only two markers could be annotated using BLAST: T_595 similar to H+ transporting ATP synthase gene (Genbank FK040981.1, E-value = 7e-26), and T_1966 similar to CDA02 protein (EL696767.1, E-value= 1e-33).

All 12 InDel markers successfully amplified in all 48 samples of *C. finmarchicus* and *C. glacialis* from Greenland. The length of the PCR product for each marker was species specific, and species identification based on InDels was in agreement with morphological and molecular (mtDNA and microsatellite) identifications. No hybrids were found in our samples.

Sequences of the 16S rDNA region of *C. hyperboreus*, *C. helgolandicus* and *C. marshallae* confirmed their morphological identification (GeneBank: KF956848-KF956853).

Most of the InDel loci also amplified in these three species, and the results of cross species amplification are presented in table 1. The following loci were diagnostic: G_155 and T_461 for *C. hyperboreus*, T_3133 for *C. helgolandicus* and *C. marshallae.*

We were able to multiplex six InDel markers in a single PCR, allowing efficient discrimination of all five species. Amplification was performed in a single PCR reaction in a total volume of 5 µl with 3 ng of DNA, 1X AccuStart™ II PCR ToughMix (Quanta BioSciences) and the following combination of markers and primer concentrations (equal forward and reverse): G_150 (13 µM/µl), T_461 (27 µM/µl), T_1338 (22 µM/µl), T_1966 (16 µM/µl), T_3133 (25 µM/µl) and T_4700 (22 µM/µl). Cycling protocol was shortened to 40 minutes in total and included an initial denaturation step at 94 °C (2 min) followed by 35 cycles of 94 °C (10 s), 55 °C (10 s), 72 °C (10 s), and a final extension at 72 °C for 5 min.

Genotyping of 389 individuals from both East and West Greenland provided reliable species ID (Nielsen *et al.* 2014). In contradiction with previous studies, no hybrids between *C. glacialis* and *C. finmarchicus* were found as all individuals were assigned to either species with high confidence.

Table 1. *Calanus* InDel markers.

Marker	Primers	Amplicon size					Accession number
		C. fin.	*C. gla.*	*C. hel.*	*C. hyp.*	*C. mars.*	
G_150	F: GACGCCATTGACCATCCAGT R: GCTCCAGCGGTTAGGTTTCT	131	161	na	na	161	KF913026, KF913030
G_155	F: AGAACAACTTGAGCTTATGGA R: CAGCACAATCTTCACATTCA	188	161	170, 188	169	161	KF913027, KF913031
G_461	F: CCTGTTGCTTCAAGGTCAAA R: CTCAGGTGGATCAACCCCC	166	157	166	166	157,170	KF913028, KF913032
G_701	F: GTGGACATAGTTTACTGAAAA R: GTGAGAATGTGAGTAGAGGGCA	209	180, 191	na	na	180, 191	KF913029, KF913033
T_461	F: TGTGAAATGGCGGCCTAACA R: ACAGTACATTAAAATTATGAGCTCGCA	143	136	na	170	136	KF913034, KF913042
T_595	F: GACTTCTCCGTGAGCTCTCC R: ACCGATGACAGAGACAACCTG	68	65	68	na	65	KF913035, KF913043
T_1301	F: CCTCCTCACCCAACTCATCC R: GGATTCATATATTCAAACAAGATATCC	66	76	na	na	76	KF913036, KF913044
T_1338	F: GACTACTCCACCATCTCCGAC R: AAGACTACGGCATGTGTTG	75	72	na	na	72	KF913037, KF913045
T_1966	F: GCCATCGCTCACAAGATCCA R: CCCTTGCTTCCTGGGACATAG	102	99	102	102	99	KF913038, KF913046
T_3133	F: ACGTGAATCTCCTGTGATTTCTGA R: TTTCAAATCAGTCGAAAGCCGT	115	107	110	na	126	KF913039, KF913047
T_4700	F: TGAGGAGGAACACGTACAAGG R: TGTAGATCACCGTGCCATTGG	70	67	70	na	67	KF913040, KF913048
T_6474	F: CAAGCGCTCTCCCTCAAGAT R: GGAGATTTAGACCTGGATCTGGAT	97	91	86, 91	na	91	KF913041, KF913049

na – no amplification, G = genomic origin and T = transcriptome origin. The first accession number is given for *C. finmarchicus* sequence; the second is for *C. glacialis* sequence.

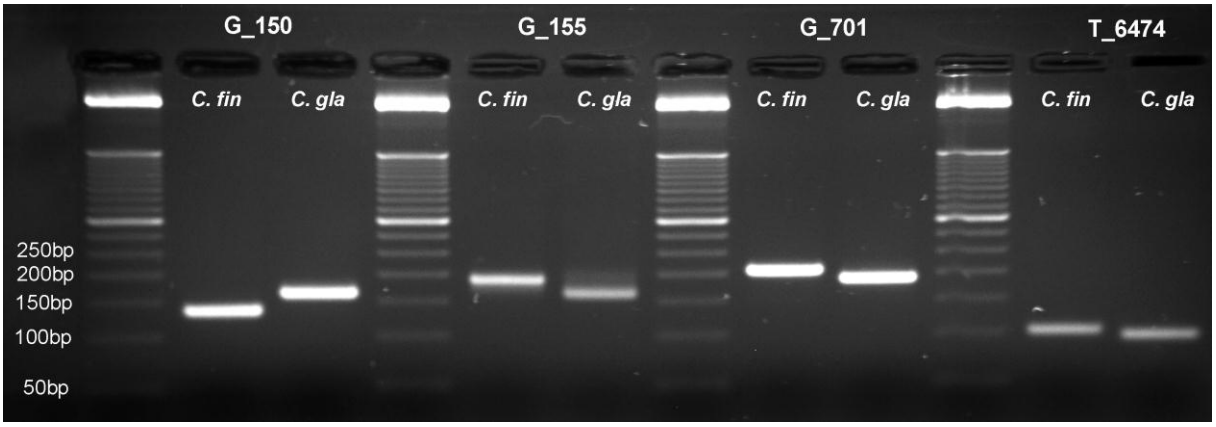

Fig 1 *Calanus* species identification using a 2.7% agarose gel and 4 InDels markers

Discussion

In less than a decade, next-generation sequencing (NGS) technologies have fundamentally changed our approach to the genomes of non-model species. NGS approaches considerably reduce the per-base sequencing cost, while dramatically increasing the number of bases sequenced, by sequencing DNA in a massively parallel fashion (Metzker 2009). Initially reserved to large scale projects, such as whole human genomes, the recent introduction of a new generation of instruments (Ion Torrent, MiSeq, 454 Junior) has opened up the use of NGS to smaller and cheaper projects (Quail *et al*. 2012). Among the new instruments, the major benefits of the Ion Torrent platform are sequencing speed and low upfront and operating costs (Perkel 2011). Indeed, sequencing costs on the Ion Torrent are <1000 USD for 1 Gb (Quail *et al*. 2012) with run time < 1/2 day. It is now economically feasible for individual laboratories to employ genome/transcriptome sequencing of non-model organisms in order to develop relevant molecular markers.

Our approach shows that both genome and transcriptome sequencing on the Ion Torrent can be used as useful tools for the development of molecular markers for copepods. Although we did not formally quantify the presence of InDels, they appeared to be more frequent and longer in the genomic contigs compared to the transcriptome based ones. However, it was generally easier to develop primers for the transcriptome based markers and amplification was also more successful. If resources were limited we would thus recommend relying on transcriptome sequencing. One of the advantages of genome sequencing is that only a small amount of DNA is required for library preparation (ca. 100 ng) compared to the amount of total RNA required (ca. 2000 ng) for transcriptome sequencing. Furthermore tissue preservation for RNA extraction can be more difficult than for DNA.

The panel of 12 nuclear markers we have developed allows the identification of all species of *Calanus* present in the North Atlantic/Arctic Oceans. The markers are easy to amplify and genotyping can be done in relatively high throughput using an automatic sequencer. Furthermore, the absence of stutter and the size difference between species allows reliable automatic scoring using software such as GENEMAPPER or GENEIOUS. For some of the markers, the size difference is sufficient for species discrimination using agarose gels (Fig. 1). This will permit the use of these markers in a low tech/low cost setting such as in the field or on board of a research vessel.

The large number of diagnostic markers between *C. finmarchicus* and *C. glacialis* we have developed will contribute to the in-depth study of hybridization between the two taxa. Hybrids between *C. finmarchicus* and *C. glacialis* have been reported in Canada (Parent *et al.* 2012), but we did not detect any in our samples from Greenland. However, the current microsatellite loci available for *Calanus* have insufficient power to separate the different classes of hybrids (*e.g.* F1, F2, backcrosses). The twelve fully diagnostic markers we have developed would allow the resolution of the different hybrid classes with low error (Anderson & Thompson 2002).

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217-1229.

Beaugrand G, Reid F, Ibañez F *et al.* (2002) Reorganization of North Atlantic marine copepod biodiversity and climate. *Science*, **296**, 1692-1694.

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. *Progress in Oceanography*, **111**, 75–90.

Bucklin A, Frost B, Kocher T (1995) Molecular systematics of six *Calanus* and three *Metridia* species (Calanoida, Copepoda). *Marine Biology*, **121**, 655-664.

Chambers G, MacAvoy E (2000) Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B*, **126**, 455-476.

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 Journal of Marine Science* 71, 241-253 doi:10.1093/icesjms/fst147

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 Molecular Biology Reporter*, **22**, 339-346.

Daase M, Vik J, Bagoien E *et al.* (2007) The influence of advection on *Calanus* near Svalbard: statistical relations between salinity, temperature and copepod abundance. *Journal of Plankton Research*, **29**, 903-911.

Everett M, Grau E, Seeb J (2011) Short reads and nonmodel species: exploring the complexities of next-generation sequence assembly and SNP discovery in the absence of a reference genome. *Molecular Ecology Resources*, **11**, 93-108.

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

Falk-Petersen S, Mayzaud P, Kattner G *et al.* (2009) Lipids and life strategy of Arctic *Calanus*. *Marine Biology Resources*, **5**, 18-39.

Gabrielsen T, Merkel B, Søreide J *et al.* (2012) Potential misidentifications of two climate indicator species of the marine arctic ecosystem: *Calanus glacialis* and *C. finmarchicus*. *Polar Biology*, **35**, 1621-1628.

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

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

Hirche H, Hagen W, Mumm N *et al.* (1994) The northeast water Polynya, Greenland Sea. 3. Mesozooplankton and macrozooplankton distribution and production of dominant herbivorous copepods during spring. *Polar Biology*, **14**, 491-503.

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

Hunter M, Hart K (2013) Rapid microsatellite marker development using next generation pyrosequencing to inform invasive burmese python—*Python molurus bivittatus*—management. *International Journal of Molecular Sciences*, **14**, 4793-4804.

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics*, **23**, 1289-1291.

Kwasniewski S, Hop H, Falk-Petersen S *et al.* (2003) Distribution of *Calanus* species in Kongsfjorden, a glacial fjord in Svalbard. *Journal of Plankton Research*, **25**, 1-20.

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

Lindeque P, Hay S, Heath M *et al.* (2006) Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four *Calanus* congeners in the North Atlantic. *Journal of Plankton Research*, **28**, 221-238.

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

Lønne O, Gulliksen B (1989) Size, age and diet of polar cod, *Boreogadus Saida* (Lepechin 1773), in ice covered waters. *Polar Biology*, **9**, 187-191.

Maar M, Møller E, Gürkan Z *et al.* (2013) Sensitivity of *Calanus* spp. copepods to environmental changes in the North Sea using life-stage structured models. *Progress in Oceanography*, **111**, 24-37.

Metzker ML (2009) Sequencing technologies — the next generation. *Nature Reviews Genetics*, **11**, 31-46.

Nagy L, Kocsube S, Csanadi Z *et al.* (2012) Re-mind the gap! Insertion – deletion data reveal neglected phylogenetic potential of the nuclear ribosomal internal transcribed spacer (ITS) of fungi. *PloS ONE*, **7**, e49794.

Nie X, Wang L, Liu P *et al.* (2012) Development of chromosome-arm-specific microsatellite markers in *Triticum aestivum* (Poaceae) using NGS technology. *American Journal of Botany*, e369-e371.

Nussberger B, Greminger M, Grossen C *et al.* (2013) Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources*, **13**, 447-460.

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

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

Pepinski W, Abreu-Glowacka M, Koralewska-Kordel M *et al.* (2013) Population genetics of 30 INDELs in populations of Poland and Taiwan. *Molecular Biology Reports*, **40**, 4333-4338.

Perkel J (2011) Making contact with sequencing's fourth generation. *BioTechniques*, **50**, 93-95.

Pompanon F, Bonin A, Bellemain E *et al.* (2005) Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics*, **6**, 847-859.

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

Provan J, Beatty G, Maggs C *et al.* (2007) Expressed sequence tag-derived microsatellites for the cool-water marine copepod *Calanus finmarchicus*. *Molecular Ecology Notes*, **7**, 1369-1371.

Quail MA, Smith M, Coupland P *et al.* (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, **13**, 341.

Selkoe KA, Toonen R (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615-629.

Søreide J, Falk-Petersen S, Hegseth E *et al.* (2008) Seasonal feeding strategies of *Calanus* in the high arctic Svalbard region. *Deep- Sea Research Part II Topical Studies in Oceanography*, **55**, 2225-2244.

Untergrasser A, Cutcutache I, Koressaar T *et al.* (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research*, **40**, e115.

Väli Ü, Brandström M, Johansson M *et al.* (2008) Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. *BMC Genetics*, **9**, 8.

Wassmann P, Duarte C, Agusti S *et al.* (2011) Footprints of climate change in the Arctic marine ecosystem. *Global Change Biology*, **17**, 1235-1249

Part 2 Population genetic structure of *Calanus sp.*

Introduction

Calanus finmarchicus has been the subject of numerous studies examining its population genetic diversity and structure, although – rather perplexingly – the results of the different studies have been inconsistent. Using mtDNA, Bucklin and Wiebe (1998) found low levels of diversity and concluded that the species may have experienced 75% range reduction and latitudinal displacement, giving rise to a genetic bottleneck. In contrast, Provan *et al.,* (2009) concluded that historical population sizes have been stable. At large spatial scales, significant differentiation between NE and NW Atlantic populations of *C. finmarchicus* was observed based on mitochondrial 16S rRNA sequence variation (Bucklin & Kocher, 1996). Using SNPs, Unal & Bucklin (2010) found support for large-scale differentiation, which may reflect persistent barriers to gene flow associated with entrainment in ocean gyres. However, microsatellite markers developed by Provan *et al.,* (2007) and cytochrome b sequence variation revealed no significant genetic differentiation (Provan *et al.,* 2009). The conflicting genetic analyses may result from the fact that all studies have limited sampling effort and used a small number of molecular markers. For *C. helgolandicus*, the little knowledge available suggest that some population structure in the Atlantic but the study suffers from limited sampling effort and a small number of molecular markers (only mtDNA) (Yebra *et al.* 2011). To our knowledge there has been no study on *C. glacialis* nor *C. hyperboreus* population genetic structure. Clearly, a reappraisal of *Calanus* population structure is needed, but using a larger number of markers (microsatellites and SNPs).

Materials and methods

Sampling

Most of the samples were collected during EURO-BASIN cruises. Unfortunately, for *C. hyperboreus* and *C. helgolandicus*, the numbers of individuals collected were too low to be useable for population genetics analysis. For *C. finmarchicus (Cfin),* six sampling locations were selected to cover the entire distribution range of the species and to provide temporal replications (Table 1). For *C. glacialis* (*Cgla*), 3 sampling location were selected based on the presence of enough individuals (Table 1). All samples were collected between 0-100 m to 0-200 m depth using WP2 or similar net with mesh size of 200 um. Animals were preserved in 95 % undenatured ethanol with subsequent change of ethanol after 24 h. Geographic distances between sample locations were calculated with Distance Calculator (http://www.daftlogic.com/projects-google-maps-distance-calculator.htm).

Table 1. Position of the sampling stations

Area	Abbreviation	Collection date	Latitude, longitude	Species
West Greenland	WG	April-2012	69.2333, - 53.3833	*Cfin, Cgla*
East Greenland	EG	August-2012	74.3097, - 20.2507	*Cfin, Cgla*
Barents Sea	BAR	August-2012	70.5012, 19.9915	*Cfin*
Norwegian Sea	NOR2012	August-2012	64.669, 0.0032	*Cfin*
Norwegian Sea	NOR2013	May-2013	65.0531, -0.8637	*Cfin*
Bay St. Lawrence	STL	April-2013	47.2728, -59.8000	*Cfin*

DNA extraction and InDel species identification

Twenty females of each species were sorted from each location according to the prosome length criteria (Kwasniewski et al. 2003). DNA was extracted individually using the E.Z.N.A. Insect DNA Kit (Omega Bio-Tek) or E.Z.N.A. DNA/RNA Isolation Kit (Omega Bio-Tek) according to manufacturer's instruction. Molecular species identification was performed using multiplex of six InDels (G_150, T_461, T_1338, T_1966, T_3133 and T_4700) according to Smolina et al. (2014) (See part 1 above).

Microsatellites genotyping

All confirmed *C. finmarchicus* and *C. glacialis* individuals were genotyped with the six most informative microsatellite loci previously developed for *C. finmarchicus*: EL696609, EL585922, EH666870 (Provan *et al.* 2007); FK868270, FK670364, FK867682 (Parent *et al.* 2012) as described elsewhere (Nielsen *et al.* 2014, Smolina *et al.* 2014). Genetic differentiation (F_{ST}, Weir & Cockerham 1984) was estimated using pair-wise permutation test implemented in Genetix v 4.05 (Belkhir et al 2001). Significance of F_{ST} values was assessed after Bonferroni correction for multiple comparisons.

SNPs genotyping by sequencing

Whole genome SNPs genotyping of *C. finmarchicus* was performed using double digest RAD (ddRAD) (Peterson et al. 2012) on a MiSeq Illumina sequencing platform at University of Nordland. Preparation of sequencing libraries was done according to Peterson et al. (2012) with minor modifications. Sixteen individuals of pure *C. finmarchicus* were pooled in equal amounts per sample location. Six sample pools of approximately 100 ng were individually digested overnight at 37^{o}C with 20000 units of *EcoRI*-HF (New England BioLabs) and SbfI-HF (New England BioLabs) enzymes each in CutSmart buffer (New England BioLabs) and total volume of 50 ul. Reactions were cleaned with the Agencourt AMPure XP system (Beckman Coulter) using 1.5x volume of the reagent. Ligation of digested DNA fragments with P1 and P2 adapters (Table 2) was performed using Quick Ligation Kit (New England BioLabs) for 10 min at room temperature, and was followed with Agencourt AMPure XP clean-up (Beckman Coulter) as described above. Ligated fragments in the range 500 – 600 bp were selected separately for each library using 2% agarose gel E-Gel SizeSelect (Invitrogen). All obtained DNA was used for a PCR amplification for 15 cycles in total volume of 50 ul with Phusion High-Fidelity PCR kit (New England BioLabs) according to manufacturer´s instruction and annealing temperature of 62 ^{o}C. Reactions were cleaned with the Agencourt AMPure XP system (Beckman Coulter) using 0.8x volume of the reagent, size selected using 2% agarose gel E-Gel SizeSelect (Invitrogen) and run on the Agilent 2200 TapeStation System (Agilent Technologies) for quantification. Libraries were pooled in equal amounts and prepared for sequencing with MiSeq Reagent Kits v2 (Illumina) on a 500 cycles chip (Illumina).

Digestion of *C. glacialis* DNA was not successful. The restriction pattern obtain was not compatible with the library preparation. This could be the result of the complexity of *C. glacialis* genome. Indeed its genome is expected to be about twice the size of C. finmarchicus, implying numerous duplication events. Such complex genomes are still a challenge for Next Generation Sequencing methods.

Table 2. Library adapters and barcodes

Adaptor	Sequence
GCATG_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCATG
GCATG_EcoRI_P1.2	/5Phos/AATTCATGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AACCA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACCA
AACCA_EcoRI_P1.2	/5Phos/AATTTGGTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
TCGAT_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCGAT
TCGAT_EcoRI_P1.2	/5Phos/AATTATCGAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
CGATC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGATC
CGATC_EcoRI_P1.2	/5Phos/AATTGATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
AAGGA_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAGGA
AAGGA_EcoRI_P1.2	/5Phos/AATTTCCTTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
ATTAC_EcoRI_P1.1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATTAC
ATTAC_EcoRI_P1.2	/5Phos/AATTGTAATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT
SbfI_P2.1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCTGCA
SbfI_P2.2	/5Phos/GGAGATCGGAAGAGCGAGAACAA

Genome-wide SNP analysis

Sequenced reads were demultiplexed according to the six barcodes using DDemux (Rasic *et al.* 2014). A special pipeline for ddRAD population genomics, dDocent (Puritz *et al*. 2014), was used for quality filtering (Phred > 20), *de novo* assembly of digest fragments and read mapping to the assembled reference contigs. Mapped bam files were merged into one file with mpileup program in SAMtools (Li *et al*. 2009) and transferred into synchronized file for PoPoolation2 (Kofler *et al*. 2011) software for SNP calling and calculation of genetic differentiation. Pair-wise F_{ST} were estimated for each SNP with coverage more than 160x (10x per individual) and more than 320x (20x per individual). Significance of pair-wise F_{ST} per SNP was assessed with Fisher´s exact test.

Results

Microsatellite-based genetic differentiation

For *C. finmarchicus*, all six microsatellites were used for the estimation of genetic differentiation, while in *C. glacialis*, only three loci were polymorphic and thus used for the analysis: EL696609, EH666870 and FK868270. Pair-wise F_{ST} between all populations in *C. finmarchicus* (Table 3) and *C. glacialis* (Table 4) were low and not significant. No correlation between F_{ST} and geographic was detected.

Genetic diversity (heterozygosity) of populations within species was comparable between location within species but contrasting between the two species, with values of 0.49 for *C. finmarchicus* and 0.12 for *C. glacialis,* respectively (Table 5).

Table 3. Pair-wise F_{ST} based on six microsatellites (right uppermost) and coastal/ocean distance (left lowermost, km) between *C. finmarchicus* populations.

F_{ST}	WG	EG	BAR	NOR2012	NOR2013	STL
WG	-	0.02	0	-0.01	0	0
EG	3326	-	0	0.04	0.01	0.01
BAR	4395	858	-	0	-0.01	0
NOR2012	3593	811	672	-	-0.02	0
NOR2013	3548	776	670	18	-	-0.01
STL	2587	2537	3139	2482	2457	-

Table 4. Pair-wise F_{ST} based on six microsatellites (right uppermost) and coastal/ocean distance (left lowermost, km) between *C. glacialis* populations.

F_{ST}	WG	EG	SVB
WG	-	0.02	-0.00
EG	3326	-	-0.03
SVB	4565	1231	-

Table 5. Microsatellite based diversity in *C. finmarchicus* and *C. glacialis* populations.

Population	*C. finmarchicus*	*C. glacialis*
WG	0.49	0.11
EG	0.41	0.12
BAR	0.46	-
NOR2012	0.53	-
NOR2013	0.53	-
STL	0.49	-
SVB	-	0.15

SNP-based genetic differentiation

For *C. finmarchicus,* MiSeq Illumina sequencing resulted in total 4,358,899 reads, a one seventh of them was PhiX sequencing control. Number of cleaned reads for each library was as follow: 684831 (WG), 681440 (EG), 648361 (BAR), 781038 (NOR2012), 558814 (NOR2013), and 308710 (STL). *De novo* assembly resulted in 41,514 contigs. Data analysis were performed for SNPs with 160x (Table 6) and 320x coverage (Table 7). Genetic diversity estimates were similar across the range (Table 8).

Table 6. Pair-wise F_{ST} based on genome-wide SNPs (right uppermost) and number of SNPs with significant F_{ST} (left lowermost) between *C. finmarchicus* populations with 160x coverage per SNP (total 4672 SNPs).

N\F_{ST}	WG	EG	BAR	NOR2012	NOR2013	STL
WG	-	0.0095	0.0087	0.0090	0.0089	0.0094
EG	1232	-	0.0085	0.0092	0.0090	0.0094
BAR	1290	1313	-	0.0085	0.0085	0.0087
NOR2012	1076	1211	1217	-	0.0088	0.0086
NOR2013	1199	1337	1359	1173	-	0.0090
STL	839	924	904	847	884	-

Table 7. Pair-wise F_{ST} based on genome-wide SNPs (right uppermost) and number of SNPs with significant F_{ST} (left lowermost) between *C. finmarchicus* populations with 320x coverage per SNP (total 1482 SNPs).

N\F_{ST}	WG	EG	BAR	NOR2012	NOR2013	STL
WG	-	0.0062	0.0054	0.0061	0.0055	0.0061
EG	390	-	0.0056	0.0064	0.0058	0.0061
BAR	405	414	-	0.0057	0.0052	0.0057
NOR2012	365	392	398	-	0.0055	0.0063
NOR2013	390	417	402	346	-	0.0057
STL	283	288	299	301	287	-

Table 8. SNPs based diversity in *C. finmarchicus* populations.

Population	*C. finmarchicus*
WG	0.32
EG	0.31
BAR	0.30
NOR2012	0.31
NOR2013	0.30
STL	0.33

Discussion

For *Calanus finmarchicus,* based on both microsatellites and SNPs, our results indicate that there is no population structure across the entire North Atlantic. Genetic diversity is also homogeneous and relatively high in all populations. For *Calanus glacialis,* microsatellites markers also show a lack of population structure in the North Atlantic. Genetic diversity is also homogeneous but lower in comparison to *C. finmarchicus.* For both species it is thus likely that North Atlantic populations are part of single large panmictic units.

The rate and absolute magnitude of climate change in the last 150 yrs (since emission of greenhouse gases began with industrialization) is expected to be greater than that of the last four million years and, as such, can be considered a human-mediated and planet-wide experiment in unusually rapid selection (Overpeck *et al.*, 2005, Reusch & Wood, 2007). Genetic responses to climate-induced selection, therefore, will be important in mitigating the negative aspects of climate change (Gienapp *et al.*, 2008). In general, global environmental change invokes two basic responses of organisms: ecological (dispersal, phenotypic plasticity) and evolutionary (genetic change), both of which are integrated into a “move, be plastic or evolve” strategy (Jackson & Overpeck, 2000). On the short term, plasticity and shifting distribution are the most important mechanism for coping with environmental changes. For *C. finmarchicus* and *C. glacialis*, temperature stress experiment (See WP4) showed that *C. finmarchicus* appears to be better equip in term of molecular mechanisms to deal with stress in comparison to *C. glacialis* (WP4). In addition, *C. finmarchicus* from the mid-distribution range appears to be less stressed by increased temperature compared to northern individuals. Given the lack of population structure within *C. finmarchicus* across its entire ranges it is much likely the results of acclimatization and not local adaptation. Regarding range shift, the capacity of a species to track rising temperature will be depend of the level of gene flow and the connectivity of populations.

Modelling (Bryant et al. 1998; de Young et al. 2004; Speirs et al. 2006) and early population genetic studies (Bucklin & Kocher 1996; Bucklin et al. 1996) of *C. finmarchicus* have yielded somewhat conflicting results with respect to dispersal and connectivity. Recent modelling studies have suggested that there may be extensive connectivity and transport across the entire range of *C. finmarchicus* (Speirs et al. 2006). Our results are support with such prediction and also confirmed the results from earlier microsatellites studies (Provan et al. 2009).

The level of genetic variation within a population can affect its productivity, growth and stability (Hughes et al. 2008) and enhance resistance and resilience of populations to perturbations (e.g. Reusch et al. 2005, Massa et al. 2013). Genetic variation is thus considered a key to evolutionary adaptation and survival of species in the long term (Bijlsma & Loeschcke 2012) especially under climate change. We found contrasting level of genetic diversity for *C. finmarchicus* and *C. glacialis*. The lower genetic diversity found in *C. glacialis,* in combination with its apparent lack of heat shock response (see WP4) and the prediction for species distribution modelling (see Villarino et al. submitted, and WP3 D3.8, which indicates that *C. glacialis* is expected to suffer higher latitudinal shift northward (11.3º) compared to *C. finmarchicus* (3.7º)) suggests that *C. glacialis* populations in the North Atlantic are more vulnerable to climate change than *C. finmarchicus.* A decrease/loss of *C. glacialis* population could have strong impact on the Arctic marine ecosystem as *Calanus glacialis* is one of the key species in the Arctic. On the edge of the sea ice, it can account for up to 80% of zooplankton biomass (ACIA 2004). It is the main primary consumer and, due to their ability to synthesize and bio-accumulate lipids for overwintering and reproduction (Scott *et al.* 2000), older, lipid-enriched development stages are the main

food for commercially important pelagic fishes, e.g. herring (*Clupea harengus*), mackerel (*Scomber scombrus*) and capelin (*Mallotus villosus*), and for planktivorous invertebrates and birds like the little auk (*Alle alle*) (Varpe et al., 2005).

References:

ACIA (2004) In: Hassol, S. (Ed.), Impacts of a Warming Arctic: Arctic Climate Impact and Assessment. Cambridge University Press, Cambridge, 144 pp

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme, F (2001). GENETIX, logiciel sous WindowsTM pour la genetique des populations. Laboratoire Genome et populations, Montpellier, France.

Bijlsma & Loeschcke (2012). Genetic erosion impedes adaptive responses to stressful environments

Bryant A.D, Hainbucher D, Heath M (1998) Basin-scale advection and population persistence of *Calanus finmarchicus. Fish. Oceanogr.* 7, 235–244

Bucklin A, Kocher T.D (1996) Source regions for recruitment of *Calanus finmarchicus* to Georges bank: evidence from molecular population genetic analysis of mtDNA. *Deep Sea Res. II.* 43, 1665–1881.

Bucklin A, Sundt R, Dahle G (1996) Population genetics of *Calanus finmarchicus* in the North Atlantic. Proceedings of a I.C.E.S. planning workshop for a TransAtlantic Study of Calanus finmarchicus. *Ophelia* 44, 29-45.

de Young B, Heath M, Werner F, Chai F, Megrey B, Monfray P (2004) Challenges of modeling ocean basin ecosystems. *Science.* 304, 1463–1466.

Evol Appl, 5, 117–129

Gienapp P, Teplitsky C, Alho JS, Mills JA, MerilÄ J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology* 17, 167-178.

Hughes et al. (2008). Ecological consequences of genetic diversity *Ecology Letters,* 11, 609-623

Jackson ST, Overpeck JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26, 194-220.

Kofler R, Vinay Pandey R, Schloetterer C (2011) PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, 27, 3435–3436.

Kwasniewski S, Hop H, Falk-Petersen S et al. (2003) Distribution of *Calanus* species in Kongsfjorden, a glacial fjord in Svalbard. *Journal of Plankton Research*, 25, 1-20.

Li H, Handsaker B, Wysoker A et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25, 2078-2079.

Massa et al. (2013). Entangled effects of allelic and clonal (genotypic) richness in the resistance and resilience of experimental populations of the seagrass *Zostera noltii* to diatom invasion *BMC Ecology,* 13, 1-12

Nielsen TG, Kjellerup S, Smolina I et al. (2014) Live discrimination of *Calanus glacialis* and *C. finmarchicus* females: can we trust phonological differences*? Marine Biology*, 161, 1299-1306.

Overpeck J, Cole J, Bartlein P (2005) In: Climate Change and Biodiversity, (Lovejoy, T. E. & Hannah, L., eds.). pp. 91-108. Yale University Press, New Haven, CT.

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

Peterson BK, Weber JN, Kay EH et al (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7, e37135. doi:10.1371/journal.pone.0037135.

Provan J *et al.* (2009) High dispersal potential has maintained long-term population stability in the North Atlantic copepod *Calanus finmarchicus. Proc Roy Soc London B* 276:301-307

Provan J, Beatty G, Maggs C *et al.* (2007) Expressed sequence tag-derived microsatellites for the cool-water marine copepod *Calanus finmarchicus*. *Molecular Ecology Notes*, 7, 1369-1371.

Puritz J, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, 10, e431. doi: 10.7717/peerj.431.

Rašić G, Filipović I, Weeks AR et al. (2014) Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. *BMC Genomics*, 15, 275.

Reusch et al. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity *PNAS*, 102, 2826–2831

Reusch TBH, Wood TE (2007) Molecular ecology of global change. *Molecular Ecology* 16, 3973-3992.

Scott CL, Kwasniewski S, Falk-Petersen S, Sargent JR (2000) Lipids and life strategies of Calanus finmarchicus Calanus glacialis and Calanus hyperboreus in late autumn, Kongsfjorden, Svalbard. *Polar Biology* 23, 510-516.

Smolina I, Kollias S, Poortvliet M, et al. (2014) Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for *Calanus* species (Copepoda: Calanoida) identification. *Molecular Ecology Resources*, 14, 1072-1079.

Speirs D.C, Gurney W.S.C, Heath M.R, Horbelt W, Wood S.N, de Cuevas B.A (2006) Ocean-scale modeling of the distribution, abundance and seasonal dynamics of the copepod *Calanus finmarchicus. Mar. Ecol. Prog. Ser.* 313*,* 173–192

Varpe, Ø., Ø. Fiksen, and A. Slotte. (2005). Meta-ecosystems and biological energy transport from ocean to coast: the ecological importance of herring migration. *Oecologia* 146, 443–45

Weir BS & Cockerham CC (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*, 38, 1358.

Yebra L *et al.* (2011) Barriers in the pelagic: population structuring of Calanus helgolandicus and C. Euxinus in European waters *MEPS* 428: 135-149

Table S1

Designated purpose	Species	N	Location of collection	Coordinates of location	Collection date	Collected by	Morphological identification	Molecular identification	Accession number
Genome sequencing	*C. finmarchicus*	1	Norway: Svinøy Island	62.3685 N, 05.2027 E	aug-2012	I. Smolina	Prosome length	6 microsatellites	KF913026-KF913029
Genome sequencing	*C. glacialis*	1	Norway: Rijpfjorden	80.3028 N, 22.3017 E	sep-2012	I. Smolina	Prosome length	6 microsatellites	KF913030-KF913033
Transcriptome sequencing	*C. finmarchicus*	30	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites	KF913034-KF913041
Transcriptome sequencing	*C. glacialis*	30	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites	KF913042-KF913049
Marker optimization	*C. finmarchicus*	2	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites	Not applicable
Marker optimization	*C. glacialis*	2	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites	Not applicable
InDels test	*C. finmarchicus*	12	West Greenland Sea	69.2333 N, 53.3833 W	mai-2011	T. G. Nielsen	Prosome length, redness	6 microsatellites, 16S rDNA mtDNA, 12 InDels (present study)	Not applicable
InDels test	*C. glacialis*	12	West Greenland Sea	69.2333 N, 53.3833 W	mai-2011	T. G. Nielsen	Prosome length, redness	6 microsatellites, 16S rDNA mtDNA, 12 InDels (present study)	Not applicable
InDels test	*C. finmarchicus*	12	East Greenland Sea	74.3097 N, 20.2507 W	aug-2012	M.D. Agersted	Prosome length, redness	6 microsatellites, 16S rDNA mtDNA, 12 InDels (present study)	Not applicable
InDels test	*C. glacialis*	12	East Greenland Sea	74.3097 N, 20.2507 W	aug-2012	M.D. Agersted	Prosome length, redness	6 microsatellites, 16S rDNA mtDNA, 12 InDels (present study)	Not applicable
Cross-species amplification	*C. hyperboreus*	2	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	General morphology	16S rDNA mtDNA, 12 InDels (present study)	KF956850, KF956851
Cross-species amplification	*C. helgolandicus*	2	Spain: Bay of Biscay	43.7000 N, 6.1500 W	mai-2003	L. Blanco-Bercia	General morphology	16S rDNA mtDNA, 12 InDels (present study)	KF956848, KF956849
Cross-species amplification	*C. marshallae*	2	USA: Alaska	55.3722 N, 168.175 W	sep-1999	B. Frost	General morphology	16S rDNA mtDNA, 12 InDels (present study)	KF956852, KF956853
InDel sequencing	*C. finmarchicus*	2	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites, 12 InDels (present study)	KF913026-KF913029, KF913034-KF913041
InDel sequencing	*C. glacialis*	2	Greenland: Disko Bay	69.2333 N, 53.3833 W	apr-2012	E. Moller and P. Lindeque	Prosome length, redness	6 microsatellites, 12 InDels (present study)	KF913030-KF913033, KF913042-KF913049