Eastern Baltic cod (*Gadus morhua*) mtDNA diversity through time

Master's thesis

in the single-subject Master's program, subject Biological Oceanography
of the Faculty of Mathematics and Natural Sciences
of the Christian-Albrechts-University of Kiel

Submitted by

Érika Endo Kokubun

First examiner: Prof. Dr. Thorsten B. Reusch
Second examiner: Prof. Dr. Reid Brennan

Kiel, December 2022

Table of contents

Abstract

The overharvesting of fish stocks has led to populational declines, both in numbers and physical conditions. Additionally, larger individuals are under stronger pressure, as it is a commercially desirable trait and due to the size-selectivity of fishing gear. When the targeted traits have underlying genetic variability and are heritable, then exploitation could be detected via genetic changes through a process known as fisheries-induced evolution (FIE). Despite the decrease of fishing pressure, the Eastern Baltic (EB) cod has shown a decreasing size at maturity and no signs of recovery. Therefore, this study aimed to assess the temporal trends in the mitogenome diversity of EB cod and to scan for genomic evidence of FIE. From five time points (1996, 2002, 2008, 2014, 2019), 117 sequences (respectively: 31, 22, 24, 20, and 20 individuals) were compared to neighboring Atlantic cod populations (Western Baltic: 22 sequences, and North sea: 24) and a paleontological group (14 sequences), regarding genetic diversity indices, differentiation and genetic relationships. Variant annotation and allele frequency trends were also analyzed. The EB cod mitogenome has maintained stable nucleotide and haplotype diversities for the past 23 years. A weak differentiation between the Baltic populations and the outgroups was detected. Haplotype data captured signals from past expansion events, when the full haplotypic diversity migrated into the Baltic sea around 10 kya. SNPs annotation and allelic frequencies suggest that functional constraints, varying mutation rates, and a high effective population size are the strongest drivers of maintained polymorphisms within the population over the time considered, rather than fisheries pressure. Hence, mtDNA appears to not have enough power to detect short time scale changes, but it still is a useful marker for deep history information. Therefore, fisheries management could use the mitogenome to establish population baselines in order to support better management actions. Ultimately, the apparent lack of selection in the mtDNA does not exclude the impact of fisheries, but rather that stock recovery will not be hindered by mitochondrial diversity loss.

Keywords: Fisheries-induced evolution, Atlantic cod, mitogenome, Baltic sea, population genetics.

Introduction

Fisheries-induced evolution.

Humans have been exploiting wild fish populations even before modern fishing practices (Erlandson & Rick, 2010). Nevertheless, the potential role of fisheries in the evolution of wild populations was not acknowledged for a long time (Stokes & Law, 2000), despite it being a topic of discussion for decades already (Handford *et al.*, 1977; Law & Grey, 1989; Ricker, 1981). The harvest of wild populations has the potential to cause genetic changes even when it's not selective, since its intensity and prolonged duration increase mortality rates, which then selects for earlier maturation (Law, 2007; Swain & Benoît, 2015). Even so, fisheries are seldom non-selective, as it usually targets certain desired phenotypes (e.g. larger individuals) or the most available fish seasonally and/or spatially (Andersen *et al.*, 2012). If such artificial selection is capable of changing gene frequencies, it could induce adaptive evolution of the population against predation, in this case harvesting by humans, a process now known as fisheries-induced evolution (FIE) (Heino *et al.*, 2015). FIE, therefore, takes place when the specific traits that render fish more vulnerable to fisheries have a genetic basis (i.e. are heritable) and feature population-level variability (Heino *et al.*, 2013; Heino *et al.*, 2015).

According to life-history theory, higher mortality rates of adult individuals favors early maturation of the population and/or allocation of energy to reproduction rather than somatic growth (Law, 2000; Stearns *et al.*, 2000). Consequently, when the larger fish are under strong selective pressure by fisheries, the population shifts towards smaller sizes and earlier maturity due to induced changes in growth rates (Heino & Godø, 2002). For instance, Mollet *et al.* (2007) found supporting evidence that the exploitation of the North Sea sole (*Solea solea*) has been playing a role in the decrease in age and size-at-maturation of the stock. Moreover, the selection for larger fish also affects other phenotypic traits, such as average size (Bianchi *et al.*, 2000; Ricker, 1981; Uusi-Heikkilä *et al.*, 2015), physical conditions (Eero *et al.*, 2015), and diverse physiological and behavioral characteristics (Sbragaglia *et al.*, 2021; Tillotson & Quinn, 2018; Walsh *et al.*, 2006) in exploited stocks. How harvesting exactly affects wild populations is dependent on the specific life-history characteristics and harvesting regimes on the stock, therefore exact predictions should be done on a case by case basis (Enberg *et al.*, 2012; Poos *et al.*, 2011). Beyond population life-history characteristics, FIE can also lead to effects in population structure and dynamics (Kuparinen & Merilä, 2007), lower fitness in their natural environment (Heino *et al.*, 2013),

and/or even bring maladaptive traits that could hinder the recovery of the population (Walsh *et al.*, 2006).

Given that life‐history, morphological, and behavioral characteristics can be tied to genetic traits that render them heritable, it is reasonable to expect that fisheries will lead to evolutionary changes (Guerra *et al.*, 2020; Law, 2007; Petrou *et al.*, 2021; Young *et al.*, 2020). Using genetic markers from modern and historical samples of Icelandic cod (*Gadus morhua*), Jakobsdóttir *et al.* (2011) detected long-term temporal declines of the *Pan* I^{bb} genotype, coherent with changes in the patterns of exploitation which resulted in the disappearance of older cod individuals. Similarly, an overall decrease of the *vgll3*L* allele, associated with late maturity, was found in an Atlantic salmon (*Salmo salar*) population found in the Teno river, mainly driven by indirect effects of overharvesting their prey species, the capelin (*Mallotus villosus*) (Czorlich *et al.*, 2022). Similar findings have been reported for other exploited populations, such as cod (Therkildsen *et al.*, 2013), salmon (Consuegra *et al.*, 2005), snapper (Hauser *et al.*, 2002), and Mediterranean swordfish (Righi *et al.*, 2020). Genetic diversity seems to be decreasing overall in overfished populations, especially in those with smaller effective population size, low gene flow, and overlapping generations (Pinsky & Palumbi, 2014). Considering that marine fish populations commonly present an effective population size much smaller than the census size (Hauser *et al.*, 2002), even populations with large spawning stock biomasses are endangered of losing genetic diversity, adaptability, and population persistence (Frankham *et al.*, 2002).

Changes in genetic diversity have the potential to magnify consequences from overfishing and ecological challenges, as these are very slow to reverse if there is no strong pressure to counteract the harvest selection (de Roos, 2006; Loder, 2005). Thus, it becomes relevant to assess if fisheries overexploitation is leading to a loss in genetic diversity since this is the basis for adaptation and individual fitness (Frankham *et al.*, 2002; Reed & Frankham, 2003). Consequently, a decrease in genetic diversity increases the risk of extinction (Frankham, 2005; Spielman *et al.*, 2004), decreases the recovery potential of stocks, and reduces biological productivity (Olsen *et al.*, 2004; Ryman *et al.*, 1995; Walsh *et al.*, 2006). Moreover, if such losses are accompanied by unmanaged exploitation of fish with specific traits, the population risks entering a state of irreversible collapse (Kuparinen *et al.*, 2014), especially when additionally paired with ecological processes, such as Allee effects (Winter *et al.*, 2020), environmental fluctuations, demographic stochasticity, etc (Wootton & Pfister, 2013). The Allee effect in depleted fish populations, for instance, could explain the

lack of recovery of the stocks even after substantial fishing reductions, as it has been observed for a cod population off the coast of Canada (Hutchings & Reynolds, 2004).

The Eastern Baltic Cod.

The Atlantic cod (*Gadus morhua*) is one of the most economically important species for fisheries in the North Atlantic, with more than 100 years of research (Marteinsdóttir & Rose, 2019). This species became the dominant predatory fish of the continental shelf of the northern seas and currently has 26 recognized stocks, with potentially more reproductive units (Marteinsdóttir & Rose, 2019). Two of these stocks are found in the Baltic Sea: the Western Baltic (WB) cod and the Eastern Baltic (EB) cod. These populations differ from other Atlantic populations because they live in a brackish, semi-enclosed sea with a horizontal gradient of salinity that decreases towards the northeastern areas (Snoeijs-Leijonmalm *et al.*, 2017). The Baltic Sea also presents a strong vertical salinity gradient due to the surface riverine inflow (low salinity) and the bottom inflow from the North Sea (high salinity), creating hypoxic, eventually anoxic, waters at the bottom (Snoeijs-Leijonmalm *et al.*, 2017). Therefore, the Baltic cod, in particular the EB stock, has adapted to a low salinity environment prone to hypoxic conditions, and consequently developed a relationship between the unique environmental conditions and its recruitment and distribution (Righton & Metcalfe, 2019). The EB cod is considered a separate population from WB cod due to a limited geographical overlap (Neuenfeldt *et al.*, 2013; Righton & Metcalfe, 2019), in which individuals of both stocks co-occur in the Arkona basin while genetic exchange among them is rarely detected (Hemmer-Hansen *et al.*, 2019). Their genetic isolation is due partly to their reproductive isolation caused by a mismatched spawning season and partly due to the egg characteristics of the EB cod, which are locally adapted to their habitat (Nielsen *et al.*, 2001).

The EB cod was overexploited for decades until the closure of fisheries in 2019, leading to the decline in fishing mortality (ICES, 2019b). Despite that, the stock hasn't been recovering as expected, as there has been an absence of larger cod adults, decline of spawning stock biomass, continuous decline of physical conditions, and a limited range of distribution (Eero *et al.*, 2015; ICES, 2019b). Considering that the Baltic ecosystem is currently subject to a strong multi-factorial stress (Reusch *et al.*, 2018), the EB cod population is being continuously exposed to stressful conditions due to climate change, nutrient loading, regime shifts, and deoxygenation (Carstensen *et al.*, 2014; Casini *et al.*, 2009; Mackenzie *et al.*, 2007; Meier *et al.*, 2019). These environmental factors pose ecological challenges to the

population through the deterioration of its spawning grounds, the expansion of hypoxic conditions, the increased consumption of its eggs by sprat and herring, and the disruption of food webs (Casini *et al.*, 2009; Casini *et al.*, 2016; Köster *et al.*, 2005; Limburg & Casini, 2019). One aspect of special concern is the rise in egg mortality due to the expansion of anoxic conditions. The low salinity of the Baltic causes eggs to have negative buoyancy and sink; if the eggs sink into a hypoxic or anoxic zone, mortality will be severe (Chabot & Claireaux, 2019). The production of eggs with enough buoyancy to float above these areas is dependent on parental characteristics, in particular related to the larger and older females in good physical conditions (Vallin & Nissling, 2000).

It is possible that genetic diversity declines may have occurred alongside the demographic and phenotypic decreases, as reported in cod populations of the southern Gulf (Swain *et al.*, 2007). If similar losses of genetic diversity are occurring in the EB cod, it could become an additional obstacle in the stock recovery. When losing genetic diversity, the population loses adaptive potential, becoming more susceptible to extinction risks and less resilient to environmental shocks (e.g. climate change) and increased fisheries (Winter *et al.*, 2020). As one of the most economically important stocks in Baltic Sea fisheries (ICES 2019a), an assessment and monitoring of their genetic diversity is vital for fisheries management of this exploited population.

Mitochondrial DNA.

One commonly used methodology to assess genetic variation is the analysis of mitochondrial DNA (mtDNA) diversity. Animal mtDNA has been used as a genetic marker due to its inheritance in "blocks" (i.e haplotypes), high mutation-fixation rate, limited repair ability, higher gene conservation, elevated copy numbers per cell, small genome size, etc (Desalle *et al.*, 2017; Hartl & Clark, 1997; Harrison, 1989; Wan *et al.*, 2004). Moreover, the maternal inheritance of mtDNA presents another advantage in this study. The survival of eggs in future hypoxic/anoxic conditions is dependent on the presence of larger female adults (Vallin & Nissling, 2000), which are also an ideal target to fisheries due to their larger sizes. Therefore, a bottleneck effect would likely leave signatures in the mitogenome. The unique set of characteristics of the mtDNA enable a higher resolution of evolutionary and demographic processes within a few generations than other genetic markers (Desalle *et al.*, 2017; Nyström *et al.*, 2006; van der Valk *et al.*, 2018; Xenikoudakis *et al.*, 2015). In population ecology, mtDNA can be a powerful tool as it provides information on population structure and differentiation, phylogeography, demographic history, and diversity (Ganbold *et*

al., 2020; Gariboldi *et al.*, 2016; Jørgensen *et al.*, 2018; Skurikhina *et al.*, 2018; van der Valk *et al.*, 2018).

Furthermore, analysis of long-term temporal trends in mtDNA diversity presents a unique opportunity to complement the population data, since the information from single time points may be biased by long-term temporal fluctuations, life-history traits, demographics, or even by genetic purging (Hedrick & Garcia-Dorado, 2016; Lage & Kornfield, 2006; Romiguier *et al.*, 2014). Therefore, fisheries management of the EB Cod will greatly benefit from obtaining information on current and past mtDNA diversity and its temporal trends. This will help clarify if genetic variation is declining and whether fisheries play a role in it, which would eventually enable us to predict any potential extinction risks that are relevant to the future of fishing industries.

Study objectives.

Overharvesting populations has been correlated to the decay of genetic diversity (Pinsky & Palumbi, 2014). Therefore, if the EB cod is under a strong size-selective pressure from FIE, their genetic diversity might be in decline as well. Despite evidence for decreased individual length and physical conditions (ICES, 2019b), there is not enough genomic data available to support an evolutionary response in this population. Preliminary data on our sampled population indicates a decline in phenotypic traits as well, such as a decreasing trend of average length over time (Fig. 1), and the disappearance of large cod individuals in more recent years (Fig. 2). In order to assess if there are any mitochondrial DNA changes during the last 23 years, five different time points were analyzed for diversity, differentiation, and demographic history. Two neighboring populations, from the North Sea and Western Baltic, as well as an outgroup composed of paleontological specimens, were used as reference points. Additionally, single nucleotide polymorphisms (SNPs) in EB cod were used to scan for signatures of selection, through variant annotation and temporal trends analyses.

This study aims to expand the knowledge on the mitogenome of the EB cod population while also supporting the long-term monitoring programs of the Baltic Sea. By improving our knowledge on the current state of the stock, better informed management and conservation decisions can be drawn, which then can guarantee the future of the population. Moreover, in a broader sense, the analysis of long-term temporal trends also aids the process of understanding the mechanisms of FIE and how, and/or if, we can detect its signatures of selection through mitochondrial markers.

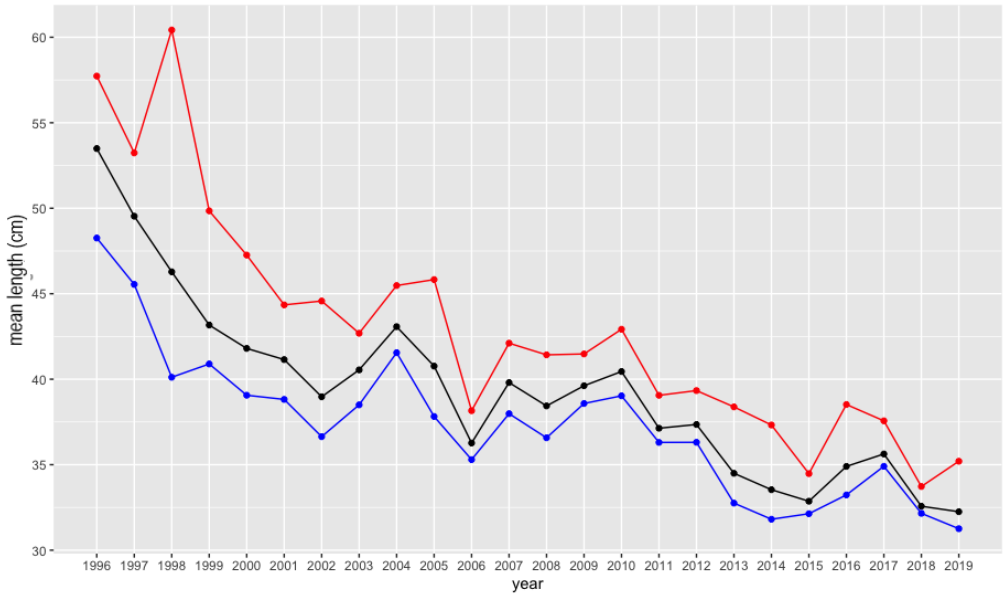

Figure 1. Mean length (cm) of sampled Eastern Baltic cod individuals caught in the Bornholm basin from 1996 to 2019. Dots indicate the mean value for the year and the connecting lines indicate trends. Black: average between males and females; red: average for females; blue: average for males.

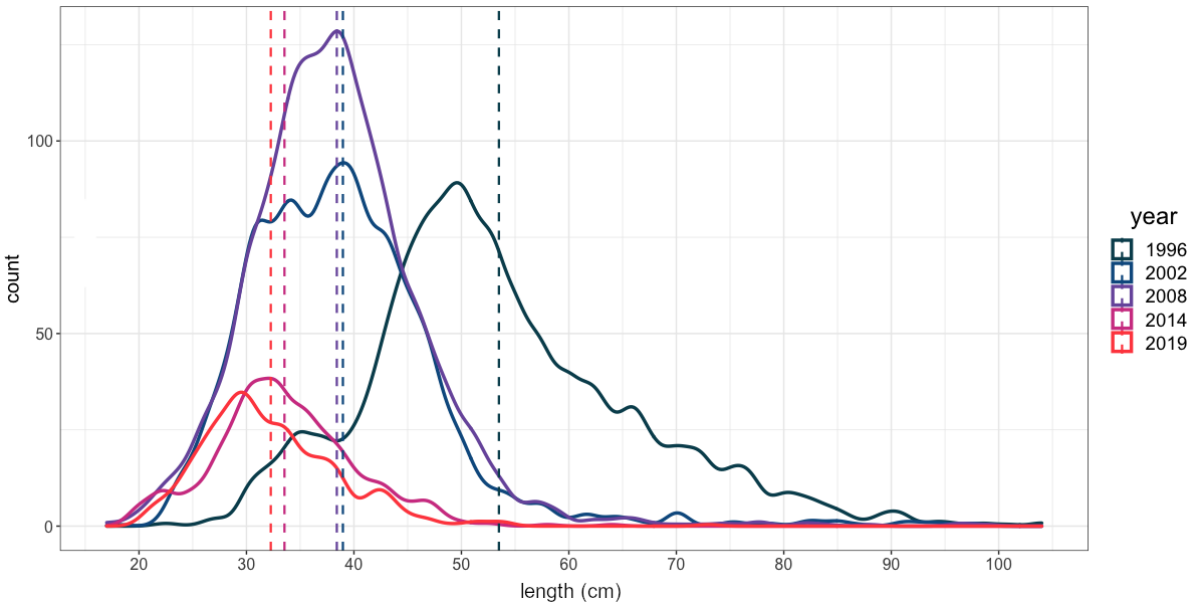

Figure 2. Count of fish individuals in length classes (cm) of sampled Eastern Baltic cod caught in Bornholm basin in the five time points selected for the study. Dashed lines indicate the mean value for each year.

Material & methods

Sampling. Genetic material was extracted from archived cod otoliths (1996 - 2012) and fin clips (2012 - 2019) collected at the annual Alkor cruise research expeditions in the Baltic Sea**.** From five time points (1996, 2002, 2008, 2014, and 2019), at least 20 mature individuals were randomly chosen to be analyzed. Within each sampling group, a balanced sex ratio was maintained.

DNA extraction and sequencing.

Otoliths. To avoid cross-contamination, any tools (e.g. forceps) utilized to manipulate the otoliths were cleaned with ethanol 70% and then sterilized in a Bunsen burner. Each otolith was incubated for around 12h in a tissue lysis solution (950 µL of lysis buffer + 50 µL of proteinase K) at 40 ºC. After incubation, the lysate solution was centrifuged at 700 RPM for 2 minutes in order to be separated from suspended solid particles (pellets). The resulting solution was separated from the pellets and centrifuged at 14.000 RCF for 3 minutes in Millipore 30 KDA columns with filters. Then, the filters were reversed in new, clean Millipore collection tubes, and centrifuged again at 1000 RCF for 2 minutes. The resulting concentrated extract was collected and purified using Qiagen QIAquick® PCR Purification Kit. The purified solution went through two more incubations at 55 ºC after adding an elution buffer (40 µL and 20 µL, before each incubation). Then, 2 µL of extracted material were analyzed in a Qubit® 2.0 Fluorometer (Thermo Fisher) in order to measure their DNA concentration. Lastly, the material was frozen at -20 ºC until sequencing, and the remaining otolith was rinsed with Milli-Q® water and stored in a new paper bag.

Microsatellite analysis. The samples went through microsatellite analysis to check for cross-contamination. First, for the PCR protocol, a master mix was prepared by mixing 400 µL of primer mix, 5 µL of each of eight primers (40 µL total), 960 µL of RNase-free water, and 500 µL of Qiagen Multiplex® PCR Master Mix. In a 96-well plate and with a stepper pipette, 9 µL of the master mix was added to each well, and then 1 µL of DNA sample was mixed in. The plate was sealed, vortexed, and centrifuged for one minute. Then, it was placed on a PCR Cycler with a cover following a pre-set program: a) initial denaturation at 96 ºC for 15 minutes; b) 27 cycles of denaturation at 94 ºC for 30 seconds + annealing at 57 ºC for three minutes + extension at 72 ºC for one minute; c) final extension at 60 ºC for 30 minutes.

When ready, in a new 96-well plate, 9 µL of a Hi-Di™ mix (one aliquot of Hi-Di™ Formamid - Thermo Fisher Applied Biosystems™ + 25 µL of GeneScan™ LIZ dye Size standard™ - Thermo Fisher Applied Biosystems™) was added to each well and mixed in with 1 µL of the PCR product. The plate was covered, centrifuged, and was put in a thermocycler for denaturation at 95 ºC for two minutes. Afterwards, the plate was inserted in the genetic analyzer (Thermo Fisher Applied Biosystems™ ABI PRISM® 3100 Genetic Analyzer) for capillary electrophoresis. Any samples presenting multiple peaks, a sign of potential cross-contamination, were excluded from further analysis.

Fin clips. For the years 2014 and 2019, DNA material was extracted from fin clips. One fin clip per individual had a small section cut (up to 25 mg) with clean, sterilized scissors, and incubated up to two hours at 56 ºC in lysis solution (180 µL of buffer ATL + 20µL of proteinase K). After the incubation period, DNA extraction was done following all the steps in the protocol Purification of Total DNA from Animal Tissues (DNeasy 96 Protocol) for the Qiagen DNeasy® Blood & Tissue Kits. The resulting material was then frozen at -20 ºC until sequencing.

Sequencing. Extracted DNA material went through whole-genome resequencing aiming for an average depth of 15x. Using Illumina DNA Prep® kit (Illumina, San Diego, USA), a pilot sequence run for 16 individuals from the 1996 population was performed by the Ancient DNA Laboratory at the Institute of Clinical Molecular Biology (IKMB) on an Illumina 6000 S4 Flowcell. Library preparation of the remaining otolith samples (15 additional otoliths from 1996; 22 otoliths from 2002; 24 otoliths from 2008) was performed at the Norwegian sequencing center using Illumina Nextera® DNA Library Preparation kit (Illumina, San Diego, USA). The finclip samples (20 finclips from 2014; 20 finclips from 2019) were prepared using the Illumina DNA Prep® kit (Illumina, San Diego, USA), and sequenced on an Illumina 6000 S4 Flowcell (Illumina, San Diego, USA) with 2x150 bp paired-end-sequencing, at the Competence Centre for Genomic Analysis (CCGA) Kiel.

Bioinformatics and variant calling. A visual workflow of how the sequences were prepared for downstream analyses can be seen in Figure 3.

Preparation of sequences. Raw sequence reads of the temporal populations were converted to BAM files using GATK (McKenna *et al.*, 2010) *FastqToSam* v. 4.0.12 command line. The

resulting unmapped sequences had their Illumina adapters marked using GATK *MarkIlluminaAdapters* v. 4.0.3. Individuals from the 1996 population which had more than one sequence run were combined into one single sequence for further analysis, using *samtools* (Li *et al.*, 2009) merge v.1.10. Sequences were then mapped against the Atlantic cod reference mitochondrial genome (GenBank: X99772.1 by Johansen & Bakke, 1996), using *bwa-mem* v. 0.7.17 (Li, 2013) and indexed with *samtools* index v.1.10. Alignment files were further processed by removing duplicate and unmapped reads, following GATK Picard v.2.27.5 and *samtools* view v.1.10 pipelines, respectively. The resulting alignment files went through quality control with GATK Picard toolkit and *qualimap* bamqc v.2.2.2a (García-Alcalde *et al.*, 2012), and were then visually inspected in the Integrative Genomic Viewer (IGV) software v.2.10.3 (Thorvaldsdóttir *et al.*, 2013).
To compare the diversity indices of EB Cod, three outside Atlantic Cod populations (henceforth referred as "outgroups") were chosen for comparison: 22 samples from the Western Baltic (WB) (Barth *et al.*, 2019), 24 from the North Sea (NS), and 14 ancient (AN) samples from Haithabu and Schleswig paleontological sites (Martínez-Garcia *et al.*, 2021). Sequences from WB were obtained from raw reads and had the same mapping and filtering steps as previously done for EB. Samples from NS and AN were obtained already processed and filtered, and were then treated to be compatible with our EB samples: first, to standardize headers and read group names in BAM files, a samtools *view* and *sed* pipeline were used, followed by indexing with samtools *index*. Then, all non-mitochondrial reads were filtered out with bamtools *split* v.2.5.1 (Barnett *et al.*, 2011) followed by visual inspection with samtools *idxstats* v.1.18.3.

Variant calling. Variant calling was performed simultaneously with all temporal and outgroup populations using GATK *HaplotypeCaller* v.4.1.4.1, with ploidy set to 1, and gVCF files were created through GATK *GenomicsDBImport* v.4.0.10.0 and *GenotypeGVCFs* v.4.0.1.2. Single nucleotide polymorphism (SNPs) variants were then filtered with *bcftools* query v.1.15 (Danecek *et al.*
For each individual in the EB and WB populations, consensus sequences were built using *bcftools* consensus v.1.15, aligned using MAFFT online tool (Katoh *et al.*, 2019), and visually inspected in Aliview v. 1.0 software (Larsson, 2014).

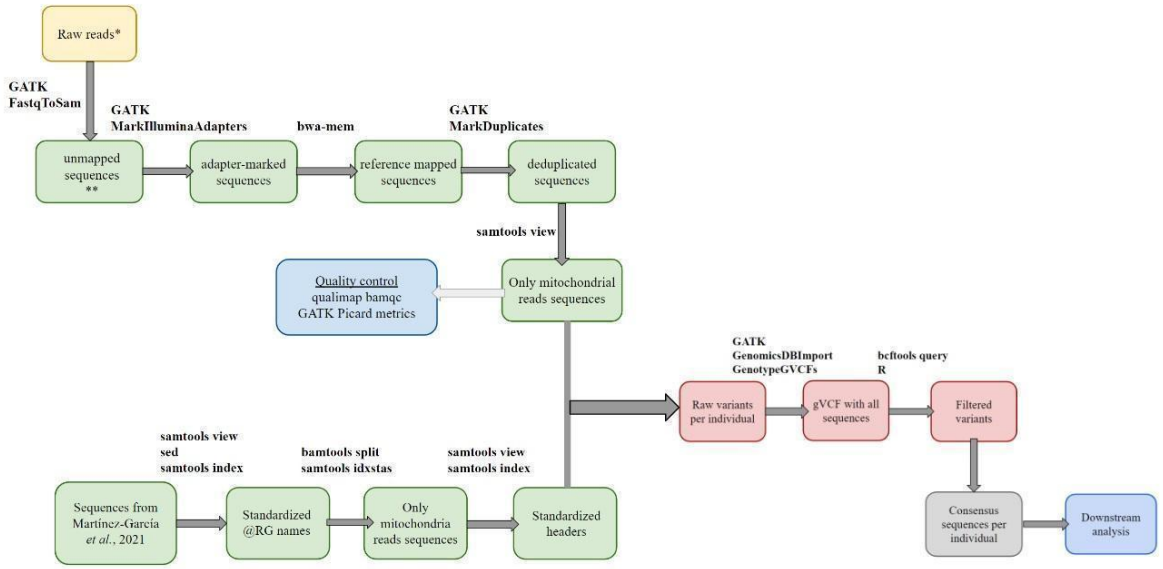

Figure 3. Bioinformatics pipeline used to prepare the sequences for downstream analyses. Bold letters indicate which toolkit or software was used. Colors indicate the type of file: yellow = fastq; green = BAM/SAM; red = VCF/gVCF; grey = FASTA; blue = diverse. * and ** indicate the starting point for EB and WB cod, respectively.

Mitochondrial genome diversity in Atlantic cod.

Mitochondrial genetic diversity. In DnaSP v.6 (Rozas *et al.*, 2017), SNPs and Indels were used to measure the genetic diversity in each population through the following diversity indices: number of haplotypes (*H*) and haplotype diversity (*Hd*), nucleotide diversity (π), number of polymorphic sites with (*S*) and without (*S**) singletons, and pairwise nucleotide differences (*k*). Additionally, to determine whether the variations are under neutral evolution, the neutrality tests Tajima's D (*TD*) ($p < 0.05$) and Fu's Fs (*F*) ($p < 0.02$) were calculated, also in Dnasp v.6.

Population differentiation. Genetic differentiation between populations were measured through Wright's F statistic (F_{ST}) and interpopulation nucleotide diversity (δ_{ST}). The dissimilarity between populations based on the values of and was then visualized through a Principal Component Analysis (PCoA) plot created in R x64 4.1.2. Additionally, pairwise genetic distances (π_{PW}) between temporal populations and outgroups (100 permutations,

1000 permutations for Mantel test) were calculated using Arlequin 3.5 (Excoffier & Lischer, 2010).

Haplotype analysis. To assess the genetic relationships between populations, haplotype networks were generated in PopArt v. 1.7 (Leigh & Bryant, 2015), using a median-joining network algorithm. Two networks were generated: one for the whole mitogenome of temporal and outgroup populations, and one only with EB haplotypes for better visualization.

Mitochondrial genome diversity in Eastern Baltic cod.

Genetic differentiation. To assess variation between different time points, an analysis of molecular variance (AMOVA) with 1000 permutations was done in Arlequin 3.5.

Variant annotation. In order to determine whether the SNPs from EB cod are located in coding (CDS) or non-coding (NC) regions, an annotation file for the Atlantic cod mitogenome was generated using MITOS web server (Bernt *et al.*, 2013). Variant annotation was then performed using *bedtools* intersect v.2.30.0 (Quinlan, 2014). The total number of occurrences of SNPs per gene was standardized as the number of variants per 100 bp and then plotted with R x64 4.1.2. Then, allele frequencies (AF) of variants were plotted per year, divided between CDS and NC regions, also in R x64 4.1.2.

Results

Sequencing and variant calling. A total of 117 samples were obtained from the temporal populations: 31 sequences for 1996 (female:male ratio of 1:0.72); 22 for 2002 (1:1.2); 24 for 2008 (1:1.18); 20 for 2014 (1:1); 20 for 2019 (1:1.2). Phred-scaled mean mapping quality was 59.6 with a mean mapping coverage of 157X. The 22 individuals from the WB population (1:1.2) had a mean mapping quality of 59.87, and a mean coverage of 206X. Variant calling after filtering, for all 177 individuals from the temporal and outside populations, found 949 variants: 914 SNPs and 35 indels. When considering only the EB cod populations, 546 variants were found, of which 466 are SNPs and 80 are indels. Singleton variants composed a considerable part of polymorphic sites for all populations analyzed (Table 1). Nevertheless, they were included in the genetic diversity indices analysis due to the high mapping quality (mean MQ: 59.72) and read depth (mean RD per site: 57326).

Mitochondrial genome diversity in Atlantic cod. All the populations analyzed had an expressive presence of unique haplotypes and no changes in diversity over time, despite slight differences between Baltic cod and the outgroups. Across temporal populations, a total of 114 haplotypes were found, associated with high haplotype diversity values (*Hd* = 1.00), with 2008 being slightly lower (*Hd* = 0.98) (Table 1). Outgroups also presented high haplotype diversity (*Hd* = 1.00). High levels of *Hd* are compatible with the high proportion of singleton variants (between 43 - 65% of total polymorphic sites). Nucleotide diversity was the same for WB and temporal populations ($\pi = 0.002$), with exception of the time point 2008 when it decreased slightly ($\pi = 0.001$) (Table 1). For AN and NS, it was estimated to be higher: $\pi = 0.004$ and $\pi = 0.003$, respectively (Table 1). When analyzed by pairwise nucleotide differences (*k*) the same pattern is observed, in which the 2008 population presents the lowest value (*k* = 25.27), while AN has the highest (*k* = 62.10) (Table 1).

Tests of neutrality indicated an excess of rare alleles in all populations. Tajima's D (TD) yielded negative values for all populations analyzed, although only 2008 (TD = -1.81; p = 0.017) and NS (TD = -1.85; p = 0.015) have statistically significant results (Table 1). Similarly, Fu's Fs was negative for all populations analyzed, but only 1996 (F = -9.108; p = 0.003) and 2002 (F = -5.10; p = 0.015) were statistically significant (Table 1).

Pairwise comparisons of genetic differentiation δ_{ST} and F_{ST} indicated low differentiation overall. Values of δ_{ST} ranged between 0.00002 and 0.00017, and the higher values were found mostly in comparisons between the AN population and 1996 (δ_{ST} = 0.00012); 2008 (δ_{ST} = 0.00015); 2014 (δ_{ST} = 0.00016); and NS (δ_{ST} = 0.00012). According to F_{ST} values, there is a very weak genetic structuring between the populations, ranging between 0.00 and 0.09. Lack of genetic structure (F_{ST} = 0.00) concentrated mostly in pairwise comparisons between temporal populations, but not for those including 2008, where some structuring was detected (0.01 - 0.04). The difference is slightly more expressive between 2008 and AN (F_{ST} = 0.05), and NS (F_{ST} = 0.09). When analyzing the PCoA results, populations from 1996, 2002, and 2019 have low dissimilarity, for both indices (Figure 4). Populations with higher dissimilarity are 2008, AN, and NS, for both indices (Figure 4).

	1996	2002	2008	2014	2019	AN	WB	NS
N	31	22	24	20	20	14	22	24
SR	1:0.72	1:1.2	1:1.18	1:1	1:1	N/A	1:1.2	N/A
H	31	22	21	20	20	14	22	24
Hd	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00
S	244	176	179	201	176	303	192	343
S*	111	101	67	87	84	107	72	119
π	0.002	0.002	0.001	0.002	0.002	0.004	0.002	0.003
k	41.82	38.6	25.27	43.87	37.17	62.10	34.20	49.76
TD	-1.22	-0.82	-1.81*	-0.94	-1.04	-1.57	-1.43	-1.85*
F	-9.108**	-5.10**	-2.82	-3.78	-4.38	-1.08	-5.66	-4.83

Table 1. Estimates of genetic diversity statistics for temporal populations of the Eastern Baltic cod (1996 to 2019) and outgroups (AN: ancient; WB: Western Baltic; NS: North Sea), using SNPs and Indels. N: sample size; SR: sex ratio (female:male); H: number of haplotypes; Hd: haplotype diversity; S: number of polymorphic sites; S*: number of polymorphic sites excluding singletons; π: nucleotide diversity; k: average number of pairwise nucleotide differences; TD: Tajima's D; F: Fu's Fs. Significant values of neutrality test are marked as * ($p < 0.05$) and ** ($p < 0.02$). Sex ratio for AN and NS were not available (N/A).

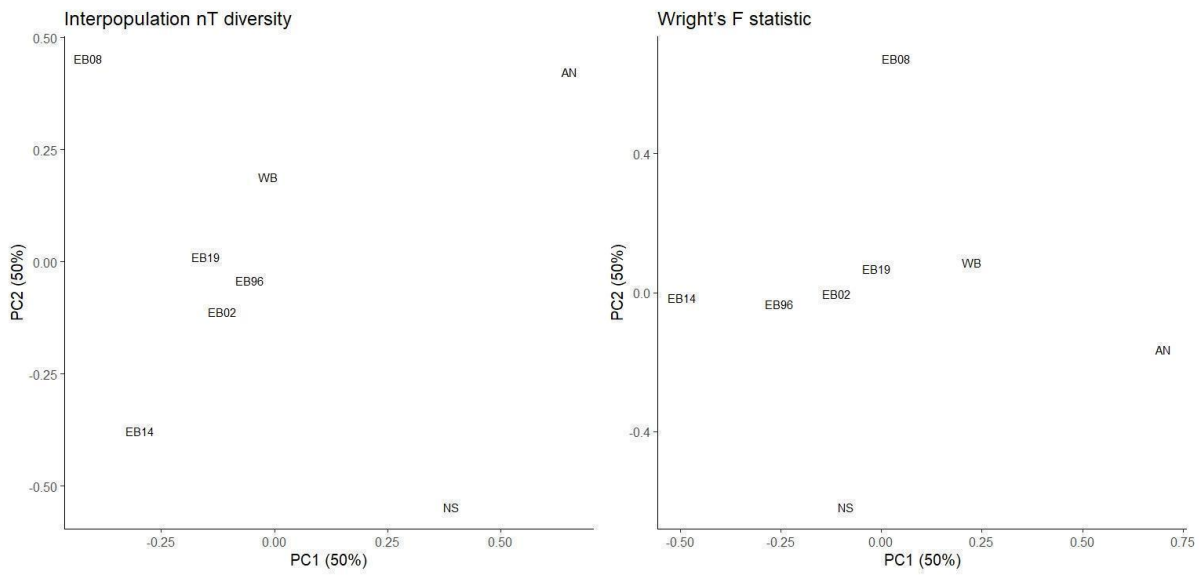

Figure 4. Genetic differentiation between temporal populations of the Eastern Baltic cod (1996: EB96; 2002: EB02; 2008: EB08; 2014: EB14; 2019: EB19) and outgroups (AN: ancient; WB: Western Baltic; NS: North Sea). Left: interpopulation nucleotide diversity (δ_{ST}); right: Wright's F statistic (F_{ST}).

Population pairwise nucleotide differences (π_{PW}) indicate that temporal populations do not differ significantly amongst themselves (except between 1996 and 2008) and WB cod. The variation is higher when compared to the outgroups. The differences were higher between AN and all other analyzed populations, while 2008 and NS shared the most similar values (Table 2). Statistically significant comparisons (p-value < 0.05) were found between AN and 1996 (2.82), 2008 (3.39), 2014 (3.15), and NS (1.87); between 2008 and 1996 (1.6), and NS (3.99).

The haplotype genealogy does not show a strong differentiation between populations, neither temporally or spatially (Fig. 5 & 6). A limited structuring is seen between Baltic samples (both WB and EB) and NS and AN samples (Fig. 5). Haplotypes from the outgroups AN and NS are majorly concentrated in between clusters, including one exclusive middle branch, but with no shared haplotypes. The only shared haplotypes are those from the Baltic populations, including WB. Even when considering only temporal populations, there is limited structuring (Fig. 6). Almost all nodes are composed of a single haplotype, indicating

the presence of many unique haplotypes. Exceptionally, there are three nodes with shared haplotypes of specimens from 2008, 2002, and 2014; 2002 and 2008; and 1996 and 2008.

Populations	**1996**	**2002**	**2008**	**2014**	**2019**	**AN**	**WB**	**NS**
1996	61.63	49.80	47.68	53.31	49.73	66.16	49.67	58.18
2002	0.00	38.51	35.67	40.64	38.25	54.01	37.74	46.30
2008	1.60*	1.16	30.52	38.40	35.65	51.18	34.45	45.27
2014	0.65	0.00	1.30	43.68	41.80	57.52	41.42	48.62
2019	0.00	0.00	0.60	0.17	39.58	54.31	37.58	46.96
AN	2.82*	2.22	3.39*	3.15*	1.99	65.05	52.65	60.42
WB	0.14	0.00	0.48	0.86	0.00	1.40	37.44	45.99
NS	1.34	1.02	3.99*	0.76	1.15	1.87*	1.25	52.04

Table 2. Population pairwise differences (π_{PW}) between temporal populations of the Eastern Baltic cod (1996 to 2019) and outside populations (AN: ancient; WB: Western Baltic; NS: North Sea). Diagonal elements: average differences within populations; above diagonal elements: average differences between populations; below diagonal elements: standardized average differences. Significant p-values are marked with * (p-value < 0.05).

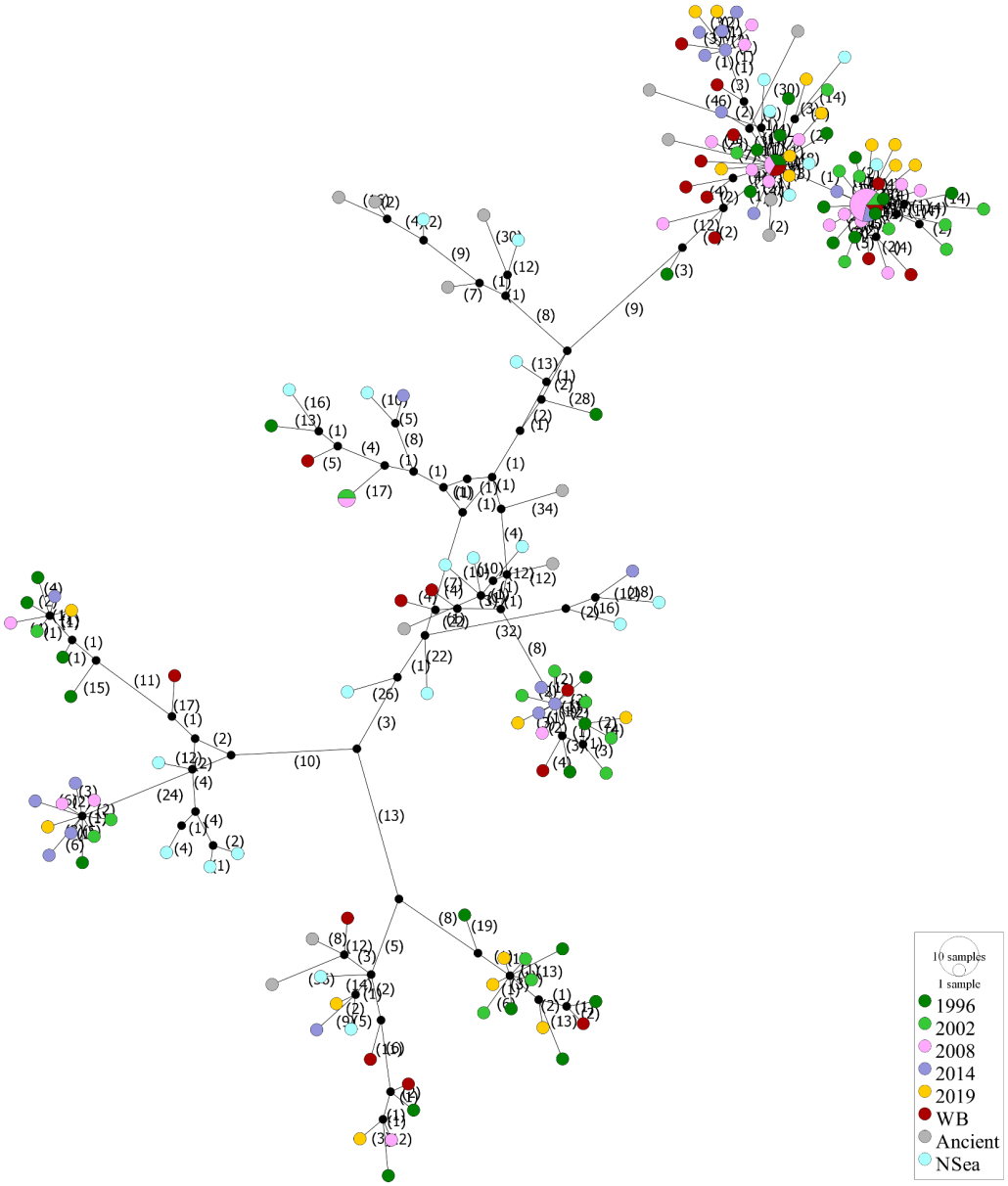

Figure 5. Haplotype median-joining network of temporal populations of Eastern Baltic cod (dark green: 1996; light green: 2002; pink: 2008; lilac: 2014; yellow: 2019) and outgroups (red: Western Baltic (WB); grey: Ancient; cyan: North Sea (N Sea)). Black dots represent median consensus sequences in between clusters. Mutation steps are shown in parenthesis.

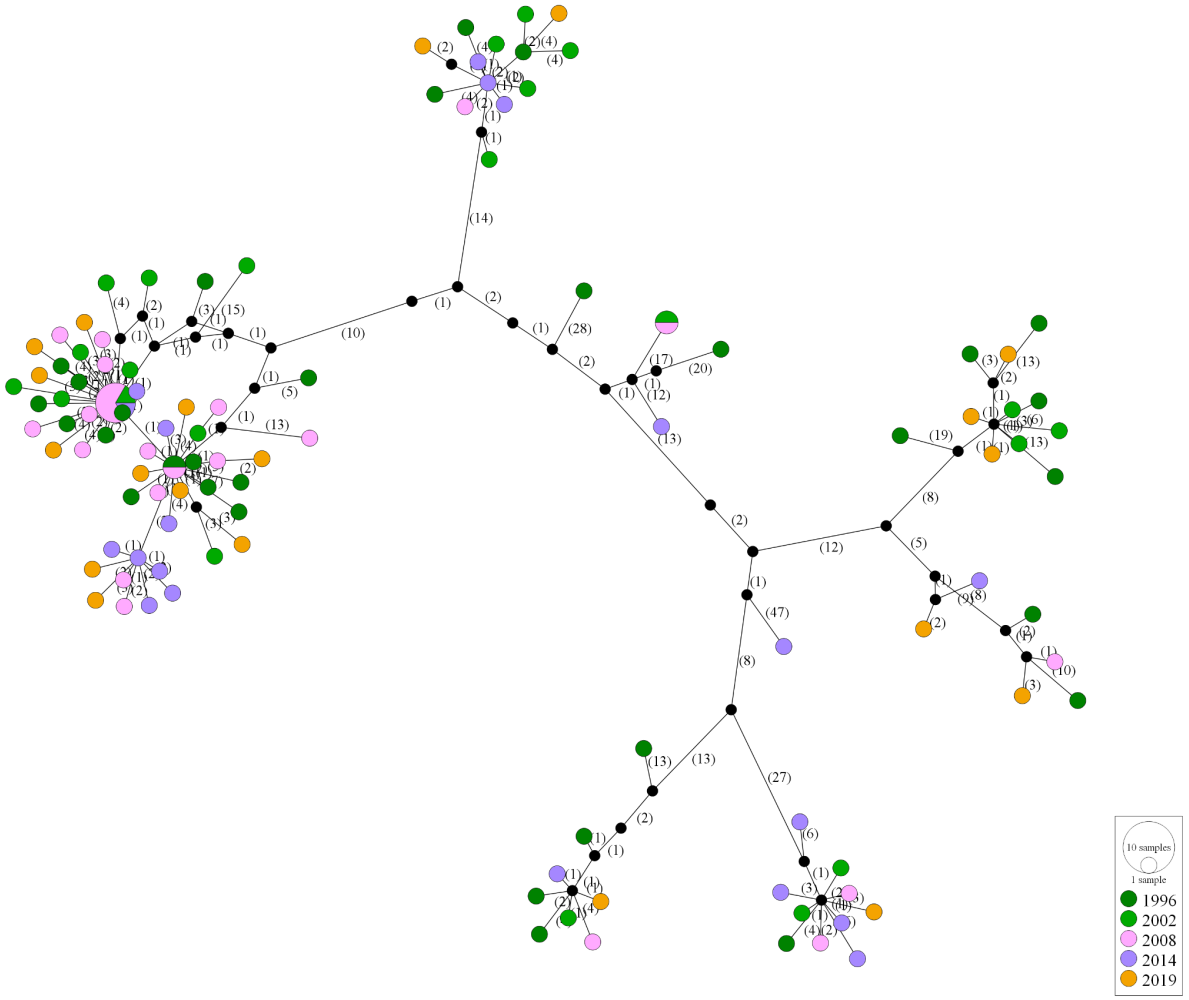

Figure 6. Haplotype median-joining network of temporal populations of Eastern Baltic cod (dark green: 1996; light green: 2002; pink: 2008; lilac: 2014; yellow: 2019). Black dots represent median consensus sequences in between clusters. Mutation steps are shown in parenthesis.

Mitochondrial genome diversity in Eastern Baltic cod. There were no detectable changes in genetic diversity over time in the EB cod population. AMOVA results (Table 3) indicated that the largest source of variation is within populations (99.37%) rather than between populations of different years (0.63%). Such lack of genetic differentiation between years is also supported by a low fixation index F_{ST} = 0.006. Accordingly, AMOVA results were not statistically significant (p-value > 0.05).

Source of variation	df	Sum of squares	Variance	% of variation
Among populations	4	101.26	0.12	0.63
Within populations	112	2470.76	22.06	99.37
Total	116	2572.02	22.20	100
Fixation Index	F_{ST}: 0.006			

Table 3. Analysis of Molecular Variance (AMOVA) between and within temporal populations of the Eastern Baltic cod (p-value > 0.05). df: degrees of freedom.

To verify whether specific genomic regions are under selective pressure, the location of the SNPs in the mtDNA was analyzed. The occurrence of variants was higher in certain regions, although no significant pattern indicative of selection was found. SNPs annotation of EB cod populations resulted in 354 variants located in CDS regions and the remaining 112 in the NC region. SNPs sites in CDS regions were distributed mostly at the genes NADH2 (44), NAD4 (40), ND5 (76), and CYTB (43), while ATP8 (2) and NAD4L (3) presented the lowest distributions (Figure 7). This distribution pattern was consistent throughout the years. In NC regions, the majority of SNPs were found in the Control region (CR) (30), 16S-rRNA (26), and 12S-rRNA (22) (Figure 8).

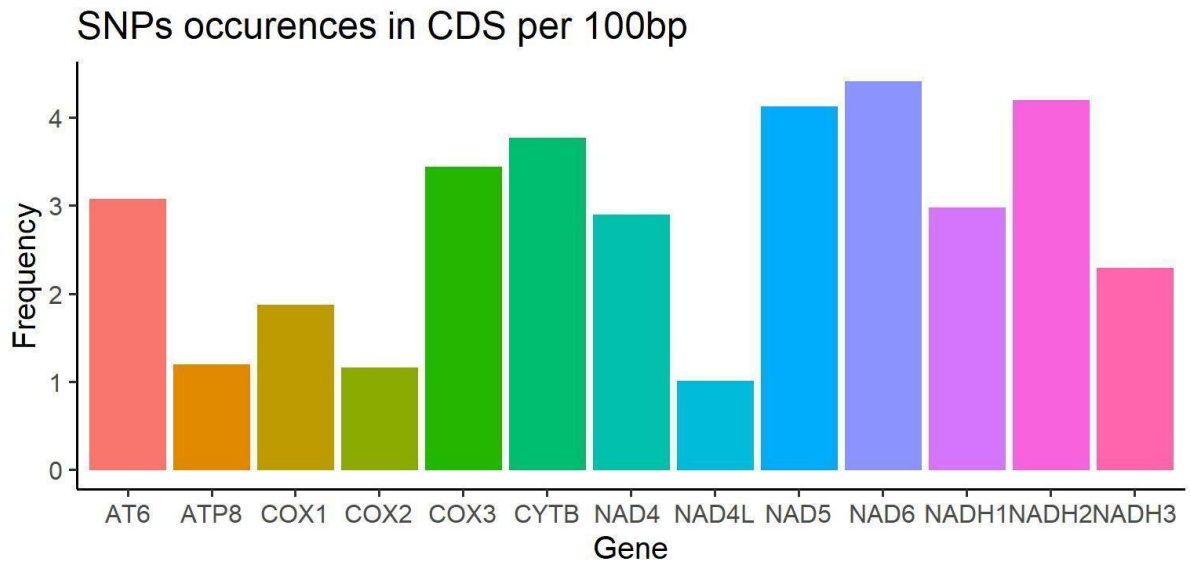

Figure 7. SNPs frequencies per coding (CDS) region in all temporal Eastern Baltic cod populations. Number of variants is standardized per 100 bp.

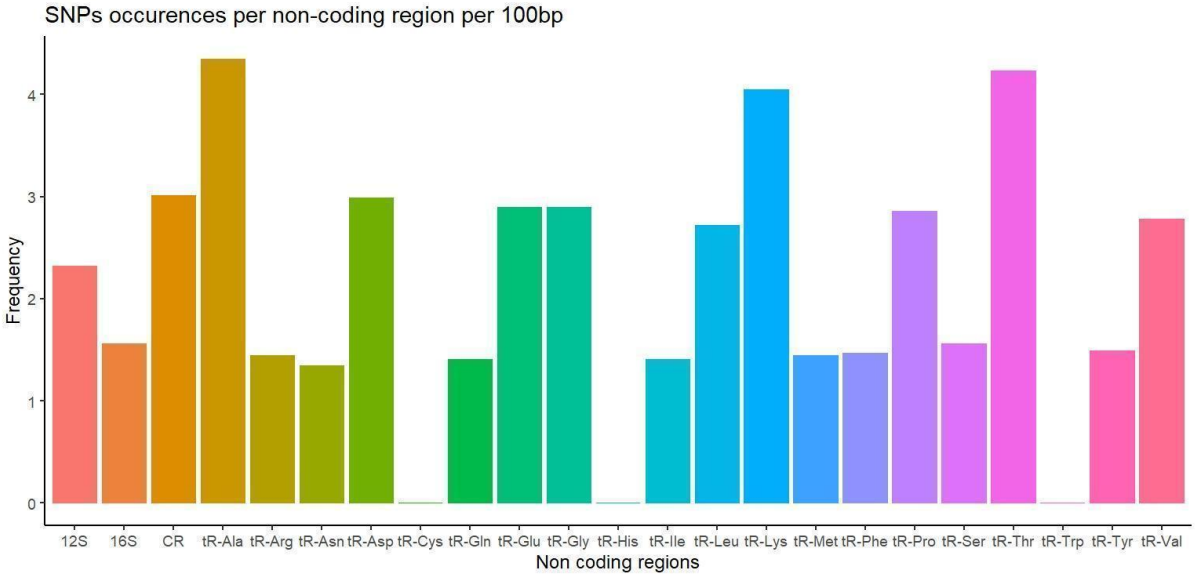

Figure 8. SNPs frequencies per non-coding (NC) region in all temporal Eastern Baltic cod populations. Number of variants is standardized per 100 bp.

An additional analysis of temporal trends of allele frequencies was done in order to verify if specific loci are under long term selective pressures. Allele frequencies of SNPs showed a few fixed variants, both in CDS (Figure 9) and NC (Figure 10) regions. In CDS, fixed alleles were mostly in genes associated with the first protein complex of the respiratory chain (NADH: ubiquinone oxidoreductase, or Complex I); the Cytochrome c oxidase (Complex IV); the Cytochrome b region; and with ATP6. In NC, fixed alleles were found in the regions of the 16S-rRNA; the tRNA-Gln; and the tRNA-Glu. Additionally, allele frequencies of some SNPs showed a distinct pattern of sudden changes between 2002 and 2008, followed by a return to previous values from 2014 onwards. In CDS, such changes were found mostly in Complex I, followed by Cytochrome b and ATP6. In NC, the changes are found mostly in the CR and intergenic spaces.

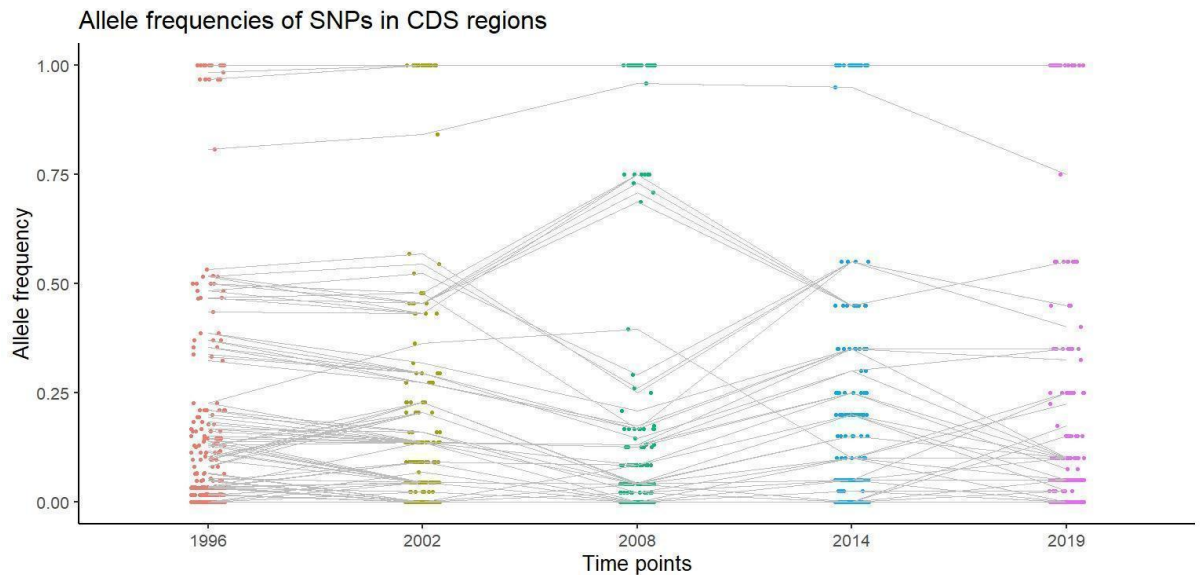

Figure 9. Temporal trends of allele frequencies of SNPs located in coding (CDS) regions of Eastern Baltic cod temporal populations (red: 1996; yellow: 2002; green: 2008; blue: 2014; pink: 2019).

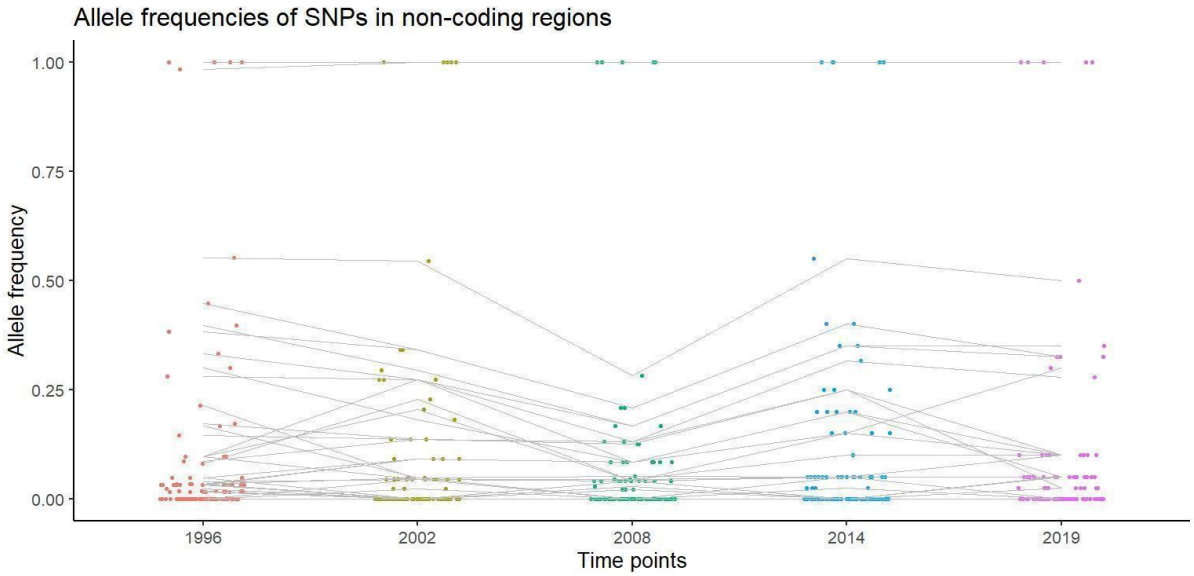

Figure 10. Temporal trends of allele frequencies of SNPs located in non-coding (NC) regions (excluding intergenic variants) of Eastern Baltic cod temporal populations (red: 1996; yellow: 2002; green: 2008; blue: 2014; pink: 2019).

To verify that the sudden allele frequency changes in 2008 were not caused by sequencing artifacts and/or lower data quality, a new quality control check was done only for the consensus sequences from 2008. Phred-score mapping quality had a minimum of 48.99 and an average of 59.92 (Figure 11), while read depth had a minimum of 6604X and a mean of 39434X (Figure 12). Given that all measures are within the threshold of high confidence calls, the data was kept in the study.

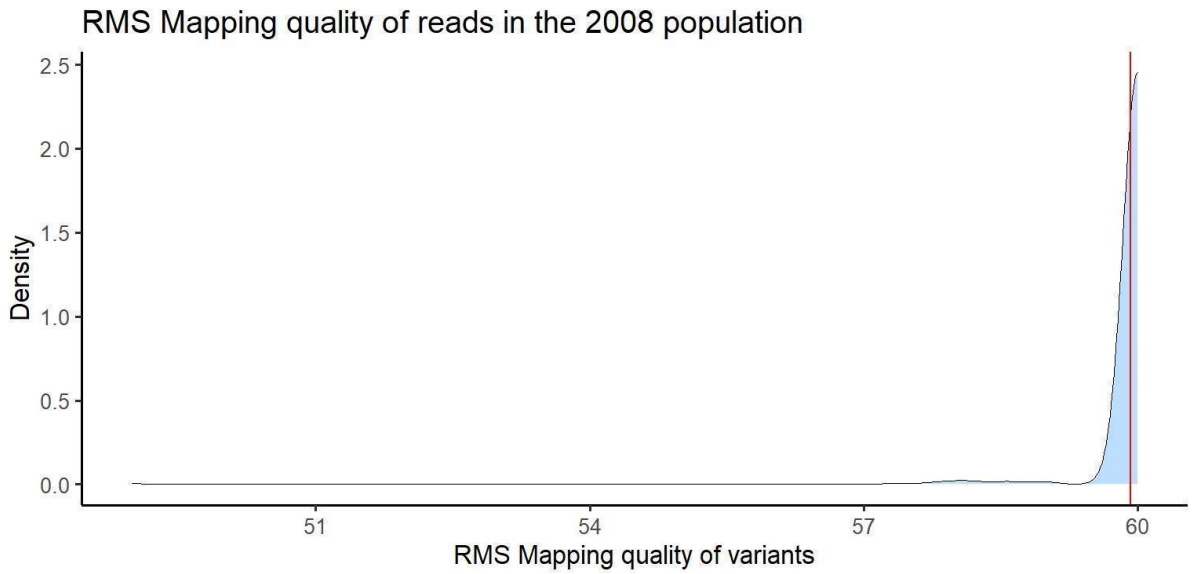

Figure 11. Density distribution of RMS (root mean square) Phred-score mapping quality of all 2008 consensus sequences. Mean (red line): 59.92.

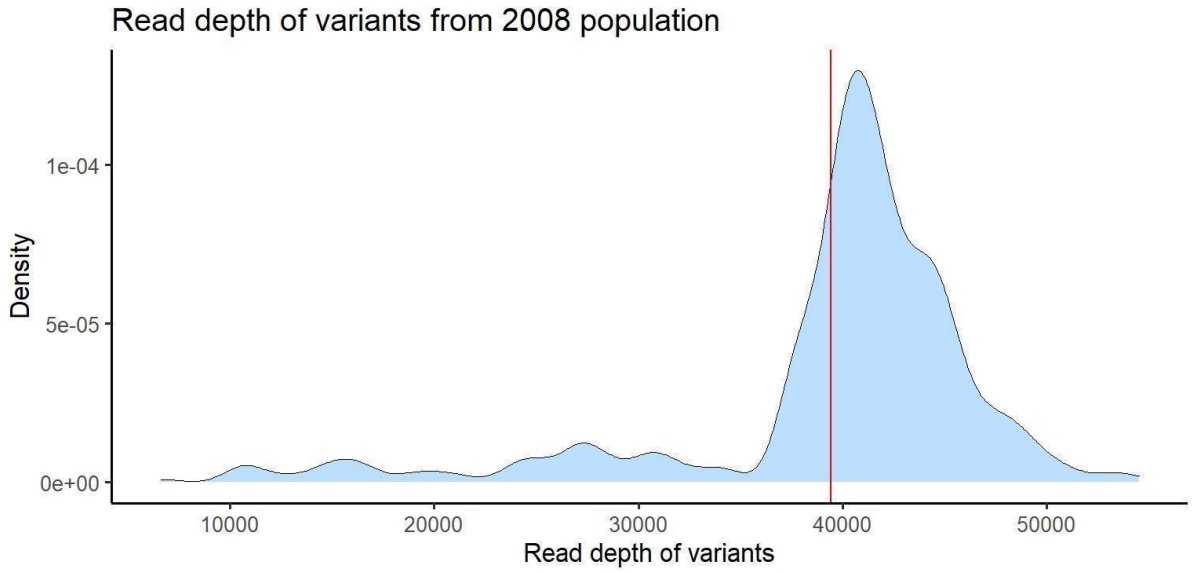

Figure 12. Read depth of SNPs for all 2008 consensus sequences. Mean (red line): 39433X.

Discussion

Genetic diversity and demographic history.

Here, temporal data of Eastern Baltic cod was compared to three outgroups of Atlantic cod in order to verify whether fishing pressures have impacted mitochondrial DNA diversity over the last 23 years. The EB cod has maintained stable mtDNA diversity levels between 1996 and 2019, with no significant changes in either of the analyzed indices. Temporal stability of genetic diversity despite intense harvesting has already been reported in nuclear markers and at the genome level of Atlantic cod, suggesting that wild populations are responding to overfishing through phenotypic plasticity and/or by maintaining high effective population sizes (N_e) (Pinsky *et al.*, 2021; Therkildsen *et al.*, 2010). The N_e of fish stocks can remain high despite declines in biomass (Poulsen *et al.*, 2005; Therkildsen *et al.*, 2010), given that it can take hundreds of generations for the decrease to be reflected in N_e (Hauser & Carvalho, 2008). Consequently, the effects could take even longer to be detected in the genome of species with high N_e, especially when considering only a single locus, such as the mtDNA (Karlsen *et al.*, 2014; Kersten *et al.*, 2021). The stability of mitogenomic diversity in exploited stocks is also seen in other populations with high N_e estimates, for instance in the Pacific herring (Moss *et al.*, 2016) and the blue jack mackerel (Moreira *et al.*, 2019).

Furthermore, the EB cod diversity indices are comparable to those found in the Western Baltic population, yet slightly lower when compared to the North Sea and ancient specimens. The higher variation in paleontological samples, nevertheless, has to be carefully interpreted, as the higher diversity levels might be an artifact caused by a mix of populations in this sub-sample. The decrease in diversity between the North sea and the Baltic, however, has been described for many populations, especially when considering mtDNA as a marker (Johannesson & Andre, 2006). Ecologically marginal environments, such as the Baltic Sea, often house populations that are peripheral to their species' main distribution (Johannesson & Andre, 2006). Genetic diversity assessments in peripheral populations have shown that these represent lower genetic variation compared to their main population (David & Wright, 2017; Schwartz *et al.*, 2003) and tend to reveal some level of differentiation (Yamaguchi *et al.*, 2010). In support of that, population differentiation analyses indicate the presence of a weak structuring between all populations, in particular separating the Baltic populations from the AN and NS groups. Although most marine fish populations have been considered to be in a

state of panmixia due to a lack of physical barriers and high dispersal potential (Craig *et al.*, 2007; Roberts & Ayre, 2010; Pan *et al.*, 2020), increasing evidence supporting local adaptation despite gene flow have challenged this assumption (Limborg *et al.*, 2012; Milano *et al.*, 2014; Nielsen *et al.*, 2009; Therkildsen *et al.*, 2013). Johannesson & Andre (2006) have found that the major driver of differentiation between Baltic and Atlantic populations is the selection of adaptive traits favorable in the unique conditions of the Baltic. For instance, the salinity and oxygen conditions have driven the adaptation of reproductive traits in EB cod (Andersen *et al.*, 2009; Kijewska *et al.*, 2016), resulting in the production of larger and lighter eggs that are neutrally buoyant in the low salinity waters of the eastern Baltic, and therefore preventing them from sinking into the hypoxic layers (Vallin & Nissling, 2000). As a consequence, divergent selection acting on reproductive traits and spawning season has risen as an efficient barrier to gene flow. Such adaptations have already been linked to genetic divergences in the genome of Baltic cod (Barth *et al.*, 2017; Berg *at al.*, 2015; Nielsen *et al.*, 2009; Wenne *et al.*, 2020), and are regarded as evidence that local adaptation is driving an ongoing process of differentiation (Berg *at al.*, 2015). Although this conclusion supports other findings of weak population structuring in mtDNA between EB, WB and NS (Martinez-García *et al.*, 2021; Nielsen *et al.*, 2005; Wenne *et al.*, 2020), the signals of population structuring are stronger in the nuclear genome (Árnason & Halldórsdóttir, 2015; Berg *et al.*, 2016; O'Leary *et al.*, 2007). This disparity suggests that the driver behind the variation in mitochondria is not the same as in the nuclear genome.

In order to elucidate whether this variation is evolving neutrally or under selective forces, the neutrality tests Tajima's D (TD) and Fu's Fs (F) were utilized in this study. All the populations presented negative values, indicating an excess of low frequency polymorphisms, a result usually associated with selective sweeps (Fu, 1997; Tajima, 1989). A selective sweep takes place when a new advantageous polymorphism eliminates and/or reduces variation in linked neutral sites (Nielsen *et al.*, 2005), consequently lowering nucleotide diversity and increasing the frequency of rare alleles (Fu, 1997; Smith & Haigh, 1974; Tajima, 1989). However, demographic forces (e.g. bottlenecks, population expansions) also affect genetic diversity and produce similar deviations from neutrality (Fu & Li, 1993; Jensen *et al.*, 2005; Ramírez-Soriano *et al.*, 2008; Tajima, 1989; Williamson *et al.*, 2005). Previous studies have shown evidence that the mitogenome of Atlantic cod appears to be under neutral evolution (Marshall *et al.*, 2009), thus our results are most likely indicators of a population expansion. Most Atlantic cod mtDNA lineages, including the NS population, have their time of origin

around 100–150 kya, in a period associated with a series of population expansions (Martinez-Garcia *et al.*, 2021) driven by major climatic events (Bigg *et al.*, 2008; Lait *et al.*, 2018). The Baltic mitochondrial lineage has appeared around 10 kya, alongside the most recent population expansion of Atlantic cod (Martinez-Garcia *et al.*, 2021). The time of origin of this lineage is in agreement with the development of the Baltic sea and the probable origin of the Baltic cod. A passageway to the Kattegat was formed between the Ancylus Lake and the North Sea, giving rise to the Littorina Sea, the predecessor of the Baltic Sea (Björck, 1995). At this time, many marine species invaded the Littorina Sea, including the Atlantic cod (Schmölcke, 2006). Therefore, it is likely that the deviation of neutrality of the EB cod mitogenome is reflecting the expansion event that gave origin to the Baltic population. This supports mtDNA as a useful marker for deep history information, but likely can not reflect shorter time scale changes, such as the one considered in this study.

The common demographic history between the outgroups and EB cod can also be visualized through the relationship between their haplotypes. Even though haplotypes from all the populations are closely related (e.g. few mutation steps in between nodes, shallow topology), the haplotypes from the NS and AN specimens are majorly concentrated in between the Baltic clusters rather than completely mixing in. The central location of AN and NS, therefore, suggests that the full haplotypic diversity of north Atlantic cod migrated into the Baltic, followed later by a gain of variants in the Baltic populations, driving their haplotypes to the tips of the branches. Omland *et al.* (2006) have described this network topology as an intermediate stage of divergence, or "intermediate polyphyly", which has been found in other populations going through processes of divergence (Peters *et al.*, 2005, Robalo *et al.*, 2020). Populations that share a common evolutionary origin tend to carry similar haplotypes, but novel haplotypes are usually not shared (Castelloe & Templeton, 1994; Moore *et al.*, 2013; Omland *et al.*, 2006; Städler *et al.*, 2005). Only three nodes share haplotypes and these are exclusive from the Baltic populations. Each population presents a high frequency of singleton variants (from 43% to 65% of the polymorphic sites), which gives origin to many unique haplotypes and consequently increases the haplotype diversity . As a result, the network has no central or highly abundant haplotype. This shallow network topology is common in mtDNA genealogies of fish due to its hyperdiversity, mostly driven by the high mutation rates of the Control region (Robalo *et al.*, 2020).

So far, this study has gathered evidence of a stable mitogenomic diversity, weak differentiation, and signatures of past demographic events. Additionally, it has highlighted the presence of unique haplotypes in the Baltic sea, suggesting that this environment houses valuable lineages for the cod mitogenome.

Polymorphisms in the Eastern Baltic cod.

According to the AMOVA, the variation found in the EB cod temporal populations was attributed mostly to differences between individuals within the same population rather than between populations. Considering the high numbers of unique haplotypes caused by the expressive presence of singletons, it is expected that the intrapopulation variation would be high. The combination of stable genetic diversity levels with weak genetic differentiation are characteristic of populations connected through high gene flow (Ferreira *et al.*, 2017; Miller *et al.*, 2009; van de Putte *et al.*, 2012). Given that our temporal populations are reproductively connected, low to no genetic differentiation was expected.

The location in which such variations are occurring can indicate the presence of selective pressures acting upon specific traits of the genome, since analyses of specific genic regions may yield a different result than full mitogenome scans (Feutry *et al.*, 2015; Jensen *et al.*, 2018) due to varying rates of mutation, functional constraints, etc (Duchêne *et al.*, 2011; Ho & Lanfear, 2010; Subramanian *et al.*, 2009). The majority of the variation found in the EB cod mitogenome was found in coding (CDS) regions rather than non-coding (NC) regions. The genes with highest SNPs occurrences were those from the first protein complex of the respiratory chain, the Complex I: ND5, NADH2, NAD4, followed by Cytochrome B and the Cytochrome c oxidase. This is a pattern already described in the literature (Jørgensen *et al.*, 2018; Marshall *et al.*, 2009), and it has been attributed to differing mutation rates caused by the strand-asymmetric replication of the mitochondria (Marshall *et al.*, 2009). The circular mitochondrial DNA is composed of a heavy strand (H-strand) and a light strand (L-strand), based on their different guanine compositions (Tanaka & Ozawa, 1994). The replication process takes place in two parts, starting at different times from different origins of replication: one for the H-strand (O_H) and one for the L-strand (O_L). Replication starts at O_H, located within the Control region of mitochondria, displacing the parental H-strand while

the process progresses towards O_L. Once the replication fork passes through O_L, the replication of the L-strand starts, then displacing the parental L-strand (Tanaka & Ozawa, 1994). A consequence of this mode of replication is that different regions of the genome are exposed as a single-strand for different lengths of time, hence providing a potential source of mutational damage (Bielawski & Gold, 2002; Sanchez-Contreras *et al.*, 2021). The two regions exposed as single-strand for the longest periods are those immediately downstream of (CYTB, ND6, ND5, and ND4) and upstream of (ND1 and ND2). In other words, the regions with the highest frequency of polymorphisms.

In NC regions, the occurrence of polymorphisms is less frequent compared to CDS. Although occurrences seem to be equivalent when standardized to 100bp, tRNA regions presented one to zero SNPs in total, indicating this region to be highly conserved in the mtDNA, as also proposed by Jørgensen *et al.* (2018). The 12S-rRNA and 16S-rRNA genes presented intermediate frequencies of SNPs in total numbers, a trend already observed in gadidae species by Marshall *et al.* (2009). These regions are considered to be conserved in mitochondria, with slower evolutionary rates than the rest of the mitogenome (Di Finizio *et al.*, 2007; Palumbi, 1996). Amongst the NC regions, the Control region presented the highest total values of SNPs. This region has been considered a hypervariable region due to reduced constraints, which accelerates its mutation rate in comparison to the rest of mtDNA (Cui *et al.*, 2009; Brown, 1985; Zhao *et al.*, 2006). Therefore, the location and frequency of SNPs found in this study, both in NC and CDS, suggest that the polymorphisms in EB cod are linked to functional constraints of the mitochondrial DNA rather than to selective forces.

The absence of selection signatures in any specific region points towards neutral evolution in the EB cod mitogenome. A dN/dS analysis could perform better at detecting evolutionary pressures in coding genes, nevertheless the presence of sequences from only one species in this study would violate the assumptions of dN/dS, thus this analysis was not done. Nevertheless, future studies should consider incorporating sequences from closely related species of *G. morhua* in order to detect selective pressures. However, previous studies have concluded, through dN/dS, that the Atlantic cod mitogenome is under neutral evolution (Marshall *et al.*, 2009). Moreover, the same patterns of polymorphism occurrences were found as in this study, which further supports our conclusion of a lack of relevant selective forces. Therefore, variant annotation analysis of the mtDNA of EB cod does not support a hypothesis of directional selection.

The temporal analysis of allele frequency changes has shown sudden oscillations between time points. Although one could argue that either a very strong selective pressure or demographic event took place around the early 2000s, the timespan of these changes is likely too short to be reflected in the mtDNA. These sudden shifts were probably caused by other factors. For instance, temporal instability could be an effect caused by the sampling design of this study and the age-structured characteristic of the EB cod population. Multiple spawners, such as the Atlantic cod (Hutchings *et al.*, 1999), present age-structured populations due to the presence of cohorts of individuals with different birth years spawning in the same season (Gotelli, 2008). In this given situation, temporal analysis of allelic frequencies (as well as estimation of effective population size, N_e) become sensitive to the sampling design. Sampling only individuals from a given age class or treating the population as a homogeneous unit (and, by consequence, randomly sampling the whole population) skews the allelic frequencies since each cohort contributes differently to the data (Waples & Yokota, 2007). Given that we randomly sampled individuals based on their sexual maturity, the year they were caught (rather than the year they were born), and not considering an extensive sampling amongst age groups, it is possible that the sudden shifts in frequencies are being driven by a generation overlap of the cohorts rather than selective pressures. At the time this study was done, there was no data on the age of the individuals analyzed, therefore this information was not taken into account in the study design. Currently, the specimens' age data is available, thus this information may be incorporated in future studies in order to avoid the effects of generation overlap.

Another plausible explanation is the time scale used in this study. Although mtDNA is considered a fast evolving molecule (Desalle *et al.*, 2017), a 23 years time frame might not be long enough to understand long term temporal trends and to differentiate between oscillations caused by directional selection and those caused by genetic drift. Short-term changes in the strength and direction of selection are expected, and might even be the norm rather than the exception in short timescales (Bell, 2010; Kingsolver & Diamond, 2011; Therkildsen *et al.*, 2013), particularly in a dynamic and stochastic habitat such as marine environments. Thus, a fast change in selection pressures might be driving the sudden oscillations in allele frequencies observed here. However, it is also possible that the oscillations are simply due to genetic drift, particularly since the oscillating sites are within the medium frequency range ($\simeq 0.5$), and that such shifts could be within normality throughout the evolutionary history of EB cod. To disentangle the effects of selection from the effects of drift (or even demographic

events) has been a challenge for evolutionary genetics, and a problem usually approached via nuclear markers (Bassitta *et al.*, 2021; Buffalo & Coop, 2020). Then, for future FIE studies, not only a longer time window might be needed, but also the inclusion of nuclear markers data should also be considered in order to properly interpret what is driving the temporal trends of allele frequencies.

Fisheries-induced evolution in the mitogenome of Eastern Baltic cod.

Although this study has not detected signs of contemporary genetic decline caused by fishing pressure over EB cod, the use of mtDNA as a molecular marker brings information on the genetic relationships with related populations and their past demographic history. The detailed analyses of genomic regions and allele frequencies supported the hypothesis of neutral evolution in the mitogenome, indicating that the polymorphisms found are linked to functional constraints, expansion events, and genetic drift.

Despite the lack of signatures of selection driven by fisheries, the information in this study still provides important insights for FIE studies. Fisheries-induced evolution is a broad term for the microevolutionary changes a population goes through caused by overharvesting of stocks based on phenotypes tied to heritable genetic traits (Heino *et al.*, 2015) . The monitoring of genetic diversity and its temporal trends, then, becomes vital for proper management of populations under pressure, in order to avoid their total collapse. The goal of fish stocks management is to secure sustainable production long term, while also promoting socio-economic well-being (Valenzuela-Quiñonez *et al.*, 2016). In this context, genomics becomes a useful tool by providing insights of possible evolutionary forces that drive temporal-spatial heterogeneity, thus allowing proper assessment of the evolutionary potential and preservation of the evolutionary legacy of exploited stocks (Valenzuela-Quiñonez *et al.*, 2016). Yet, patterns of genetic diversity integrate effects accumulated over the evolutionary path of a species, thus it is not always easy to differentiate past demographic forces and selective pressures from current fishing pressure (Therkildsen *et al.*, 2013). Mitochondrial DNA can then be used as a marker to help elucidate such questions. For instance, due to its smaller effective population size, mtDNA has the potential to preserve signals of demographic events, in a timespan relevant to fisheries, that cannot leave footprints on the larger effective population size of nuclear markers (Eytan & Hellberg, 2010). Hence, the

information provided in this study can be integrated into baseline estimates by using pre-harvest parameters as reference points to management (Marty *et al.*, 2015).

Conclusion

This study highlights the stability of the mitogenome diversity in EB cod despite a marked decline in numbers and physical conditions of the stock throughout the past decades. Moreover, it highlights that mtDNA stores information from past demographic events, such as the expansion of Atlantic cod into the Baltic sea. From then, polymorphisms were maintained by varying mutation rates within the molecule and a high effective population size. The apparent low impact that fisheries had over mitogenomic diversity does not imply that fishing has zero impact over the population, but rather that the variation in mtDNA is more closely linked to demographic processes and functional constraints of the molecule than contemporary selective forces, such as FIE. This characteristic, however, places the mtDNA as a useful tool for management decisions by, for instance, establishing pre-harvest parameters that can be incorporated into the decision-making processes. Therefore, fisheries management and conservation decisions can still benefit from the information stored in the mitogenome. Ultimately, our evidence suggests that the recovery of the Eastern Baltic cod population will not be constrained by the loss of mitochondrial genome diversity.

Acknowledgments

I would like to thank Prof. Dr. Thorsten Reusch and doctoral student Kwi Young Han for their guidance and (lots of) patience while I worked on this thesis. I would also like to acknowledge the Evolutionary Ecology group, for answering all my questions regarding specific bioinformatics details and the nature of mitochondria. I am also very thankful to the RV Alkor scientific and permanent crew, without whom this thesis would not have been possible, and who also have taught me a lot during my time on board. Lastly, but not least, I would like to thank my colleagues from the Master's program, for all the co-working sessions and support given to me in this journey.

References

Andersen, B.S., Ulrich, C., Eigaard, O.R., & Christensen, A.S. (2012). Short-term choice behaviour in a mixed fishery: investigating métier selection in the Danish gillnet fishery. ICES Journal of Marine Science, 69(1): 131-143.

Andersen, Ø., Wetten, O.F., De Rosa, M.C., Andre, C., Carelli Alinovi, C., *et al*. (2009). Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. Proceedings of the Royal Society B: Biological Sciences, 276(1658): 833-841.

Árnason, E., & Halldórsdóttir, K. (2015). Nucleotide variation and balancing selection at the Ckma gene in Atlantic cod: analysis with multiple merger coalescent models. PeerJ, 3: e786.

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., & Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics, 27(12): 1691-1692.

Barth, J.M., Berg, P.R., Jonsson, P.R., Bonanomi, S., Corell, H., *et al*. (2017). Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Molecular ecology, 26(17): 4452-4466.

Barth, J. M., Villegas‐Ríos, D., Freitas, C., Moland, E., Star, B., *et al*. (2019). Disentangling structural genomic and behavioural barriers in a sea of connectivity. Molecular ecology, 28(6): 1394-1411.

Bassitta, M., Brown, R.P., Pérez-Cembranos, A., Pérez-Mellado, V., Castro, J.A., *et al*. (2021). Genomic signatures of drift and selection driven by predation and human pressure in an insular lizard. Scientific reports, 11(1): 1-13.

Bell, G. (2010). Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1537): 87-97.

Berg, P.R., Jentoft, S., Star, B., Ring, K.H., Knutsen, H., *et al*. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua L.*). Genome biology and evolution, 7(6): 1644-1663.

Berg, P.R., Star, B., Pampoulie, C., Sodeland, M., Barth, J. M., *et al*. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. Scientific reports, 6(1): 1-12.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., *et al*. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. Molecular phylogenetics and evolution, 69(2): 313-319.

Bianchi, G., Gislason, H., Graham, K., Hill, L., Jin, X., *et al.* (2000). Impact of fishing on size composition and diversity of demersal fish communities. ICES Journal of Marine Science, 57(3): 558-571.

Bielawski, J.P., & Gold, J.R. (2002). Mutation patterns of mitochondrial H-and L-strand DNA in closely related cyprinid fishes. Genetics, 161(4): 1589-1597.

Bigg, G.R., Cunningham, C.W., Ottersen, G., Pogson, G.H., Wadley, M.R., & Williamson, P. (2008). Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. PProceedings of the Royal Society B: Biological Sciences, 275: 163–173. doi: 10.1098/rspb.2007.1153

Björck, S. (1995). A review of the history of the Baltic Sea, 13.0-8.0 ka BP. Quaternary international, 27: 19-40.

Brown, W.M. 1985. The mitochondrial genome of animals. In: Molecular Evolutionary Genetics. Ed. MacIntyre R.J. NY: Plenum, pp. 95–130.

Buffalo, V., & Coop, G. (2020). Estimating the genome-wide contribution of selection to temporal allele frequency change. Proceedings of the National Academy of Sciences, 117(34): 20672-20680.

Carstensen, J., Andersen, J.H., Gustafsson, B.G., & Conley, D.J. (2014). Deoxygenation of the Baltic Sea during the last century. Proceedings of the National Academy of Sciences, 111(15): 5628-5633.

Casini, M., Hjelm, J., Molinero, J.C., Lövgren, J., Cardinale, M., *et al*. (2009). Trophic cascades promote threshold-like shifts in pelagic marine ecosystems. Proceedings of the National Academy of Sciences, 106(1): 197-202.

Casini, M., Käll, F., Hansson, M., Plikshs, M., Baranova, T., *et al*. (2016). Hypoxic areas, density-dependence and food limitation drive the body condition of a heavily exploited marine fish predator. Royal Society open science, 3(10): 160416.

Castelloe, J., & Templeton, A. R. (1994). Root probabilities for intraspecific gene trees under neutral coalescent theory. Molecular phylogenetics and evolution, 3(2): 102-113.

Chabot, D., & Claireaux, G. (2019). Ecophysiology. In: Atlantic cod - A bio-ecology. Rose, G.A. (ed). 27-65. Vancouver/Canada: John Wiley & Sons Ltd.

Consuegra, S., De Leaniz, C.G., Serdio, A., & Verspoor, E. (2005). Selective exploitation of early running fish may induce genetic and phenotypic changes in Atlantic salmon. Journal of Fish Biology, 67: 129-145.

Craig, M.T., Eble, J.A., Bowen, B.W., & Robertson, D.R. (2007). High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). Marine Ecology Progress Series, 334: 245-254.

Cui, Z.X., Liu, Y., Li, C.P., You, F., Chu, K.H. (2009). The complete mitochondrial genome of the large yellow croaker, *Larimichthys cracea* (Perciformes, Sciaenidae): Unusual features of its control region and the phylogenetic position of the Sciaenidae. Gene, 432: 33–43.

Czorlich, Y., Aykanat, T., Erkinaro, J., Orell, P., & Primmer, C. R. (2022). Rapid evolution in salmon life history induced by direct and indirect effects of fishing. Science, 376(6591): 420-423.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., *et al*. (2021). Twelve years of SAMtools and BCFtools. Gigascience, 10(2): giab008.

David, S.R., & Wright, J.J. (2017). Genetic variation and biogeography of the spotted gar *Lepisosteus oculatus* from core and peripheral populations. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 328(7): 596-606.

De Roos, A.M., Boukal, D.S., & Persson, L. (2006). Evolutionary regime shifts in age and size at maturation of exploited fish stocks. Proceedings of the Royal Society B: Biological Sciences, 273(1596): 1873-1880.

DeSalle, R., Schierwater, B., & Hadrys, H. (2017). MtDNA: The small workhorse of evolutionary studies. Frontiers in Bioscience, 22: 873-887.

Di Finizio, A., Guerriero, G., Russo, G.L., & Ciarcia, G. (2007). Identification of gadoid species (Pisces, Gadidae) by sequencing and PCR–RFLP analysis of mitochondrial 12S and 16S rRNA gene fragments. European Food Research and Technology, 225(3): 337-344.

Duchêne, S., Archer, F.I., Vilstrup, J., Caballero, S., & Morin, P.A. (2011). Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. PloS one, 6(11): e27138.

Eero, M., Hjelm, J., Behrens, J., Buchmann, K., Cardinale, M., *et al.* (2015). Eastern Baltic cod in distress: biological changes and challenges for stock assessment. ICES Journal of Marine Science, 72(8): 2180-2186.

Enberg, K., Jørgensen, C., Dunlop, E.S., Varpe, Ø., Boukal, D.S., *et al*. (2012). Fishing‐induced evolution of growth: Concepts, mechanisms and the empirical evidence. Marine Ecology, 33(1): 1-25.

Erlandson, J.M., & Rick, T.C. (2010). Archaeology meets marine ecology: the antiquity of maritime cultures and human impacts on marine fisheries and ecosystems. Annual Review of Marine Science, 2: 231-251.

Excoffier, L., & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564-567.

Eytan, R.I., & Hellberg, M.E. (2010). Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. Evolution: International Journal of Organic Evolution, 64(12): 3380-3397.

Ferreira, D.G., Souza-Shibatta, L., Shibatta, O.A., Sofia, S.H., Carlsson, J., *et al*. (2017). Genetic structure and diversity of migratory freshwater fish in a fragmented Neotropical river system. Reviews in Fish Biology and Fisheries, 27(1): 209-231.

Feutry, P., Kyne, P.M., Pillans, R.D., Chen, X., Marthick, J.R., *et al*. (2015). Whole mitogenome sequencing refines population structure of the Critically Endangered sawfish *Pristis pristis*. Marine Ecology Progress Series, 533: 237-244.

Frankham, R. (2005). Genetics and extinction. Biological conservation, 126(2): 131-140.

Frankham, R., Ballou, J.D., & Briscoe, D.A. (2002). Conservation genetics. In: Encyclopedia of Ecology (Fath, B., ed.). Elsevier, 2780 p.

Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147(2): 915-925.

Fu, Y.X., & Li, W.H. (1993). Statistical tests of neutrality of mutations. Genetics, 133(3): 693-709.

Ganbold, O., Lee, S.H., Paek, W.K., Munkhbayar, M., Seo, D., *et al*. (2020). Mitochondrial DNA variation and phylogeography of native Mongolian goats. Asian-Australasian journal of animal sciences, 33(6): 902.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., *et al*. (2012). Qualimap: evaluating next-generation sequencing alignment data. Bioinformatics, 28(20): 2678-2679.

Gariboldi, M C., Túnez, J.I., Failla, M., Hevia, M., Panebianco, M.V., *et al*. (2016). Patterns of population structure at microsatellite and mitochondrial DNA markers in the franciscana dolphin (*Pontoporia blainvillei*). Ecology and evolution, 6(24): 8764-8776.

Gotelli, N.J. (2008). A primer of ecology. Vol. 494. Sunderland, MA: Sinauer Associates.

Guerra, A. S., Kao, A. B., McCauley, D. J., & Berdahl, A. M. (2020). Fisheries-induced selection against schooling behaviour in marine fishes. Proceedings of the Royal Society B, 287(1935): 20201752.

Handford, P., Bell, G., & Reimchen, T. (1977). A gillnet fishery considered as an experiment in artificial selection. Journal of the Fisheries Board of Canada, 34(7): 954-961.

Harrison, R.G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends in Ecology & Evolution, 4(1): 6-11.

Hartl, D.L, & Clark A.G. (1997). Principles of population genetics. 3rd edition. Sunderland/MA: Sinauer Associates.

Hauser, L., Adcock, G.J., Smith, P.J., Ramírez, J.H.B., & Carvalho, G.R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proceedings of the National Academy of Sciences, 99(18): 11742-11747.

Hauser, L., & Carvalho, G.R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish. 9: 333–362. doi: 10.1111/j. 1467-2979.2008.00299.x

Hedrick, P.W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. Trends in ecology & evolution, 31(12): 940-952.

Heino, M., & Godø, O.R. (2002) Fisheries-induced selection pressures in the context of sustainable fisheries. Bull. Mar. Sci. 70: 639–656

Heino M., Baulier L., Boukal D.S., Ernande B., Johnston F.D., *et al.* (2013). Can fisheries-induced evolution shift reference points for fisheries management? ICES J. Mar. Sci. 70(4):707–21

Heino, M., Pauli, B.D., & Dieckmann, U. (2015). Fisheries-induced evolution. Annual review of ecology, evolution, and systematics: 46.

Hemmer‐Hansen, J., Hüssy, K., Baktoft, H., Huwer, B., Bekkevold, D., *et al*. (2019). Genetic analyses reveal complex dynamics within a marine fish management area. Evolutionary applications, 12(4): 830-844.

Ho, S.Y., & Lanfear, R. (2010). Improved characterisation of among-lineage rate variation in cetacean mitogenomes using codon-partitioned relaxed clocks. Mitochondrial DNA, 21(3-4): 138-146.

Hutchings, J.A., Bishop, T.D., & McGregor-Shaw, C.R. (1999). Spawning behaviour of Atlantic cod, *Gadus morhua*: evidence of mate competition and mate choice in a broadcast spawner. Canadian Journal of Fisheries and Aquatic Sciences, 56(1): 97-104.

Hutchings, J.A., & J. D. Reynolds. (2004). Marine fish population collapses: consequences for recovery and extinction risk. BioScience(54):297–309.

ICESa. (2019). Baltic Sea Ecoregion – Fisheries overview. ICES Advice 2019 – https://doi.org/10.17895/ices.advice.5566

ICESb. (2019). Cod (*Gadus morhua*) in subdivisions 24–32, eastern Baltic stock (eastern Baltic Sea). ICES Advice on fishing opportunities, catch, and effort - Baltic Sea ecoregion. https://doi.org/10.17895/ices.advice.4747

Jakobsdóttir, K.B., Pardoe, H., Magnússon, Á., Björnsson, H., Pampoulie, C., *et al*. (2011). Historical changes in genotypic frequencies at the Pantophysin locus in Atlantic cod (*Gadus morhua*) in Icelandic waters: evidence of fisheries‐induced selection?. Evolutionary Applications, 4(4): 562-573.

Jensen, J.D., Kim, Y., DuMont, V.B., Aquadro, C.F., & Bustamante, C.D. (2005). Distinguishing between selective sweeps and demography using DNA polymorphism data. Genetics, 170(3): 1401-1410.

Jensen, E.L., Miller, J.M., Edwards, D.L., Garrick, R.C., Tapia, W., *et al*. (2018). Temporal mitogenomics of the Galapagos giant tortoise from Pinzón reveals potential biases in population genetic inference. Journal of Heredity, 109(6): 631-640.

Johannesson, K., & Andre, C. (2006). INVITED REVIEW: Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. Molecular Ecology, 15(8): 2013-2029.

Johansen, S., & Bakke, I. (1996). The complete mitochondrial DNA sequence of Atlantic cod (*Gadus morhua)*: relevance to taxonomic studies among codfishes. Molecular Marine Biology and Biotechnology, 5(3):203-14. PMID: 8817926.

Jørgensen, T.E., Karlsen, B.O., Emblem, Å., Breines, R., Andreassen, M., *et al*. (2018). Mitochondrial genome variation of Atlantic cod. BMC research notes, 11(1): 1-7.

Karlsen, B.O., Emblem, Å., Jørgensen, T.E., Klingan, K.A., Nordeide, J.T., *et al*. (2014). Mitogenome sequence variation in migratory and stationary ecotypes of North-east Atlantic cod. Marine genomics, 15: 103-108.

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in bioinformatics, 20(4): 1160-1166.

Kersten, O., Star, B., Leigh, D.M., Anker-Nilssen, T., Strøm, H., *et al*. (2021). Complex population structure of the Atlantic puffin revealed by whole genome analyses. Communications biology, 4(1): 1-12.

Kijewska, A., Kalamarz-Kubiak, H., Arciszewski, B., Guellard, T., Petereit, C., & Wenne, R. (2016). Adaptation to salinity in Atlantic cod from different regions of the Baltic Sea. Journal of Experimental Marine Biology and Ecology, 478: 62-67.

Kingsolver, J.G., & Diamond, S.E. (2011). Phenotypic selection in natural populations: what limits directional selection?. The American Naturalist, 177(3): 346-357.

Köster, F.W., Möllmann, C., Hinrichsen, H.H., Wieland, K., Tomkiewicz, J.,*et al*. (2005). Baltic cod recruitment–the impact of climate variability on key processes. ICES Journal of marine science, 62(7): 1408-1425.

Kuparinen, A., Keith, D.M., & Hutchings, J. A. (2014). Allee effect and the uncertainty of population recovery. Conservation Biology, 28(3): 790-798.

Kuparinen, A., & Merilä, J. (2007). Detecting and managing fisheries-induced evolution. Trends in ecology & evolution, 22(12): 652-659.

Lage, C., & Kornfield, I. (2006). Reduced genetic diversity and effective population size in an endangered Atlantic salmon (*Salmo salar*) population from Maine, USA. Conservation Genetics, 7(1): 91-104.

Lait, L.A., Marshall, H.D., and Carr, S.M. (2018). Phylogeographic mitogenomics of Atlantic cod *Gadus morhua*

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large data sets. Bioinformatics30(22): 3276-3278. http://dx.doi.org/10.1093/bioinformatics/btu531

Law, R. (2000). Fishing, selection, and phenotypic evolution. ICES J Mar Sci 57:659–668

Law, R. (2007). Fisheries-induced evolution: present status and future directions. Marine Ecology Progress Series, 335: 271-277.

Law, R., & Grey, D.R. (1989). Evolution of yields from populations with age-specific cropping. Evolutionary Ecology, 3(4): 343–359. doi:10.1007/bf02285264

Leigh, J.W., & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6(9):1110–1116.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN].

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., *et al*. (2009). The sequence alignment/map format and SAMtools. Bioinformatics 25: 2078–2079. doi: 10.1093/bioinformatics/btp352

Limborg, M.T., Helyar, S.J., De Bruyn, M., Taylor, M.I., Nielsen, E.E., *et al*. (2012). Environmental selection on transcriptome‐derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). Molecular ecology, 21(15): 3686-3703.

Limburg, K.E., & Casini, M. (2019). Otolith chemistry indicates recent worsened Baltic cod condition is linked to hypoxia exposure. Biology letters, 15(12): 20190352.

Loder, N. (2005). Point of No Return: Evidence is mounting that fish populations won't necessarily recover even if overfishing stops. Fishing may be such a powerful evolutionary force that we are running up a Darwinian debt for future generations. Conservation in practice, 6(3): 124-129.

MacKenzie, B.R., Gislason, H., Möllmann, C., & Köster, F.W. (2007). Impact of 21st century climate change on the Baltic Sea fish community and fisheries. Global Change Biology, 13(7)1: 1348-1367.

Marteinsdóttir, G., & Rose, G.A. (2019). Atlantic Cod: Origin and Evolution. In: Atlantic Cod: A Bio‐Ecology. Rose, G.A. (ed). 7-25. Vancouver/Canada: John Wiley & Sons Ltd.

Marshall, H.D., Coulson, M.W., & Carr, S. M. (2009). Near neutrality, rate heterogeneity, and linkage govern mitochondrial genome evolution in Atlantic cod (*Gadus morhua*) and other gadine fish. Molecular Biology and Evolution, 26(3): 579-589.

Martínez-García, L., Ferrari, G., Oosting, T., Ballantyne, R., Van der Jagt, I., *et al*. (2021). Historical demographic processes dominate genetic variation in ancient Atlantic cod mitogenomes. Frontiers in Ecology and Evolution, 342.

Marty, L., Dieckmann, U., & Ernande, B. (2015). Fisheries‐induced neutral and adaptive evolution in exploited fish populations and consequences for their adaptive potential. Evolutionary applications, 8(1): 47-63.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., *et al*. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20: 1297–1303. doi: 10.1101/gr.107524.110

Meier, H.E.M., Eilola, K., Almroth-Rosell, E., Schimanke, S., Kniebusch, M., *et al*. (2019). Disentangling the impact of nutrient load and climate changes on Baltic Sea hypoxia and eutrophication since 1850. Climate Dynamics, 53(1): 1145-1166.

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., *et al*. (2014). Outlier SNP markers reveal fine‐scale genetic structuring across European hake populations (*Merluccius merluccius*). Molecular ecology, 23(1): 118-135.

Miller, K.J., Maynard, B.T., & Mundy, C.N. (2009). Genetic diversity and gene flow in collapsed and healthy abalone fisheries. Molecular Ecology, 18(2): 200-211.

Mollet, F.M., Kraak, S.B., & Rijnsdorp, A.D. (2007). Fisheries-induced evolutionary changes in maturation reaction norms in North Sea sole *Solea solea*. Marine Ecology Progress Series, 351: 189-199.

Moore, M., Sylla, M., Goss, L., Burugu, M.W., Sang, R., *et al*. (2013). Dual African origins of global *Aedes aegypti sl* populations revealed by mitochondrial DNA. PLoS neglected tropical diseases, 7(4): e2175.

Moreira, C., Correia, A.T., Vaz‐Pires, P., & Froufe, E. (2019). Genetic diversity and population structure of the blue jack mackerel *Trachurus picturatus* across its western distribution. Journal of Fish Biology, 94(5): 725-731.

Moss, M.L., Rodrigues, A.T., Speller, C.F., & Yang, D.Y. (2016). The historical ecology of Pacific herring: Tracing Alaska Native use of a forage fish. Journal of Archaeological Science: Reports, 8: 504-512.

Neuenfeldt, S., Righton, D., Neat, F., Wright, P. J., Svedäng, H., *et al*. (2013). Analysing migrations of Atlantic cod *Gadus morhua* in the north‐east Atlantic Ocean: then, now and the future. Journal of Fish Biology, 82(3): 741-763.

Nielsen, E.E., Hansen, M.M., Schmidt, C., Meldrup, D., & Grønkjaer, P. (2001). Population of origin of Atlantic cod. Nature, 413(6853): 272-272.

Nielsen, E.E., Hemmer-Hansen, J., Poulsen, N.A., Loeschcke, V., Moen, T., *et al*. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC evolutionary biology, 9(1): 1-11.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M.J., Clark, A.G., & Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. Genome research, 15(11): 1566-1575.

Nyström, V., Angerbjörn, A., & Dalén, L. (2006). Genetic consequences of a demographic bottleneck in the Scandinavian arctic fox. Oikos, 114(1): 84-94.

O'Leary, D.B., Coughlan, J., Dillane, E., McCarthy, T.V., & Cross, T.F. (2007). Microsatellite variation in cod *Gadus morhua* throughout its geographic range. Journal of Fish Biology, 70: 310-335.

Olsen, E.M., Heino, M., Lilly, G R., Morgan, M.J., Brattey, J.*et al*. (2004). Maturation trends indicative of rapid evolution preceded the collapse of northern cod. Nature, 428(6986): 932-935.

Omland, K.E., Baker, J.M., & Peters, J.L. (2006). Genetic signatures of intermediate divergence: population history of Old and New World Holarctic ravens (*Corvus corax*). Molecular Ecology, 15(3): 795-808.

Palumbi, S.R. (1996). What can molecular genetics contribute to marine biogeography? An urchin's tale. Journal of Experimental Marine Biology and Ecology, 203(1): 75-92.

Pan, X., Ye, Z., Xu, B., Jiang, T., Yang, J., & Tian, Y. (2020). Population connectivity in a highly migratory fish, Japanese Spanish mackerel (*Scomberomorus niphonius*), along the Chinese coast, implications from otolith chemistry. Fisheries Research, 231: 105690.

Peters, J.L., Gretes, W., & Omland, K.E. (2005). Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the 'isolation with migration' coalescent method. Molecular Ecology, 14(11): 3407-3418.

Petrou, E. L., Fuentes-Pardo, A. P., Rogers, L. A., Orobko, M., Tarpey, C., *et al*. (2021). Functional genetic diversity in an exploited marine species and its relevance to fisheries management. Proceedings of the Royal Society B, 288(1945): 20202398.

Pinsky, M.L., & Palumbi, S.R. (2014). Meta‐analysis reveals lower genetic diversity in overfished populations. Molecular ecology, 23(1): 29-39.

Pinsky, M.L., Eikeset, A.M., Helmerson, C., Bradbury, I.R., Bentzen, P., *et al*. (2021). Genomic stability through time despite decades of exploitation in cod on both sides of the Atlantic. Proceedings of the National Academy of Sciences, 118(15): e2025453118.

Poos, J.J., Brännström, Å., & Dieckmann, U. (2011). Harvest-induced maturation evolution under different life-history trade-offs and harvesting regimes. Journal of Theoretical Biology, 279(1): 102-112.

Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V., & Grønkjaer, P. (2006). Long‐term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*). Molecular Ecology, 15(2): 321-331.

Roberts, D.G., & Ayre, D.J. (2010). Panmictic population structure in the migratory marine sparid *Acanthopagrus australis* despite its close association with estuaries. Marine Ecology Progress Series, 412: 223-230.

Sanchez-Contreras, M., Sweetwyne, M.T., Kohrn, B.F., Tsantilas, K.A., Hipp, M.J., *et al*. (2021). A replication-linked mutational gradient drives somatic mutation accumulation and influences germline polymorphisms and genome composition in mitochondrial DNA. Nucleic acids research, 49(19): 11103-11118.

Quinlan, A.R. (2014). BEDTools: the Swiss‐army tool for genome feature analysis. Current protocols in bioinformatics, 47(1): 11-12.

Ramírez-Soriano, A., Ramos-Onsins, S.E., Rozas, J., Calafell, F., & Navarro, A. (2008). Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. Genetics, 179(1): 555-567.

Reed, D.H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. Conservation biology, 17(1): 230-237.

Reusch, T.B., Dierking, J., Andersson, H.C., Bonsdorff, E., Carstensen, J., *et al*. (2018). The Baltic Sea as a time machine for the future coastal ocean. Science Advances, 4(5): eaar8195.

Ricker, W.E. (1981). Changes in the average size and average age of Pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences, 38(12): 1636-1656.

Righi, T., Splendiani, A., Fioravanti, T., Casoni, E., Gioacchini, G., *et al*. (2020). Loss of Mitochondrial Genetic Diversity in Overexploited Mediterranean Swordfish (*Xiphias gladius*, 1759) Population. Diversity, 12(5): 170.

Righton, D., & Metcalfe, J. (2019). Migration. In: Atlantic Cod: Origin and Evolution. Rose, G.A. (ed). 7-25. Vancouver/Canada: John Wiley & Sons Ltd.

Robalo, J.I., Francisco, S.M., Vendrell, C., Lima, C.S., Pereira, A., *et al*. (2020). Against all odds: A tale of marine range expansion with maintenance of extremely high genetic diversity. Scientific Reports, 10(1): 1-13.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., *et al*. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature, 515(7526): 261-263.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., *et al*. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. Molecular Biology and Evolution, 34: 3299-3302. DOI: 10.1093/molbev/msx248

Ryman, N., Utter, F., & Laikre, L. (1995). Protection of intraspecific biodiversity of exploited fishes. Reviews in Fish Biology and Fisheries, 5(4): 417-446.

Sbragaglia, V., López‐Olmeda, J.F., Frigato, E., Bertolucci, C., & Arlinghaus, R. (2021). Size‐selective mortality induces evolutionary changes in group risk‐taking behaviour and the circadian system in a fish. Journal of Animal Ecology, 90(2): 387-403.

Schmölcke, U. (2006). Holocene vertebrate paleontology of the Baltic Sea area. Meyniana, 58: 191-204.

Schwartz, M.K., Mills, L.S., Ortega, Y., Ruggiero, L.F., & Allendorf, F.W. (2003). Landscape location affects genetic variation of Canada lynx (*Lynx canadensis*). Molecular Ecology, 12(7): 1807-1816.

Skurikhina, L.A., Oleinik, A.G., Kukhlevsky, A.D., Kovpak, N.E., Frolov, S.V., & Sendek, D.S. (2018). Phylogeography and demographic history of the Pacific smelt *Osmerus dentex* inferred from mitochondrial DNA variation. Polar Biology, 41(5): 877-896.

Smith, J.M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. Genetics Research, 23(1): 23–35

Snoeijs-Leijonmalm, P., Schubert, H., Radziejewska, T. (2017). Biological Oceanography of the Baltic Sea. Springer Netherlands, 683 pp.

Spielman, D., Brook, B.W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences, 101(42): 15261-15264.

Städler, T., Frye, M., Neiman, M., & Lively, C. (2005). Mitochondrial haplotypes and the New Zealand origin of clonal *European Potamopyrgus*, an invasive aquatic snail. Molecular Ecology, 14(8): 2465-2473.

Stearns, S.C., Ackermann, M., Doebeli, M., & Kaiser, M. (2000). Experimental evolution of aging, growth, and reproduction in fruitflies. Proceedings of the National Academy of Sciences, 97(7): 3309-3313.

Stokes, K., & Law, R. (2000). Fishing as an evolutionary force. Marine Ecology Progress Series, 208: 307-309.

Subramanian, S., Denver, D.R., Millar, C.D., Heupink, T., Aschrafi, A., *et al*. (2009). High mitogenomic evolutionary rates and time dependency. Trends in Genetics, 25(11): 482-486.

Swain, D. P., & Benoît, H. P. (2015). Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Marine Ecology Progress Series, 519: 165-182.

Swain, D.P., Sinclair, A.F., & Mark Hanson, J. (2007). Evolutionary response to size-selective mortality in an exploited fish population. Proceedings of the Royal Society B: Biological Sciences, 274(1613): 1015-1022.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3): 585-595.

Tanaka, M., & Ozawa, T. (1994). Strand asymmetry in human mitochondrial DNA mutations. Genomics, 22(2):, 327-335.

Therkildsen, N.O., Hemmer‐Hansen, J., Als, T.D., Swain, D.P., Morgan, M.J., *et al*. (2013). Microevolution in time and space: SNP analysis of historical DNA reveals dynamic signatures of selection in Atlantic cod. Molecular Ecology, 22(9): 2424-2440.

Therkildsen, N.O., Nielsen, E.E., Swain, D.P., & Pedersen, J.S. (2010). Large effective population size and temporal genetic stability in Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Canadian Journal of Fisheries and Aquatic Sciences, 67(10): 1585-1595.

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics, 14(2): 178-192.

Tillotson, M. D., & Quinn, T. P. (2018). Selection on the timing of migration and breeding: A neglected aspect of fishing‐induced evolution and trait change. Fish and Fisheries, 19(1), 170-181.

Uusi‐Heikkilä, S., Whiteley, A. R., Kuparinen, A., Matsumura, S., Venturelli, P. A., *et al*. (2015). The evolutionary legacy of size‐selective harvesting extends from genes to populations. Evolutionary applications, 8(6), 597-620.

Valenzuela-Quiñonez, F. (2016). How fisheries management can benefit from genomics?. Briefings in Functional Genomics, 15(5): 352-357.

Vallin, L., & Nissling, A. (2000). Maternal effects on egg size and egg buoyancy of Baltic cod, *Gadus morhua*: implications for stock structure effects on recruitment. Fisheries research, 49(1): 21-37.

van de Putte, A.P., Van Houdt, J.K.J., Maes, G.E., Hellemans, B., Collins, M.A., *et al*. (2012). High genetic diversity and connectivity in a common mesopelagic fish of the Southern

Ocean: the myctophid *Electrona antarctica*. Deep Sea Research Part II: Topical Studies in Oceanography, 59: 199-207.

van der Valk, T., Sandoval-Castellanos, E., Caillaud, D., Ngobobo, U., Binyinyi, E., *et al*. (2018). Significant loss of mitochondrial diversity within the last century due to extinction of peripheral populations in eastern gorillas. Scientific reports, 8(1): 1-10.

Walsh, M.R., Munch, S.B., Chiba, S., & Conover, D.O. (2006). Maladaptive changes in multiple traits caused by fishing: impediments to population recovery. Ecology letters, 9(2): 142-148.

Wan, Q.H., Wu, H., Fujihara, T., & Fang, S.G. (2004). Which genetic marker for which conservation genetics issue?. Electrophoresis, 25(14): 2165-2176.

Waples, R.S., & Yokota, M. (2007). Temporal estimates of effective population size in species with overlapping generations. Genetics, 175(1): 219-233.

Wenne, R., Bernaś, R., Kijewska, A., Poćwierz-Kotus, A., Strand, J., *et al.* (2020). SNP genotyping reveals substructuring in weakly differentiated populations of Atlantic cod (*Gadus morhua*) from diverse environments in the Baltic Sea. Scientific reports, 10(1): 1-15.

Williamson, S.H., Hernandez, R., Fledel-Alon, A., Zhu, L., Nielsen, R., & Bustamante, C.D. (2005). Simultaneous inference of selection and population growth from patterns of variation in the human genome. Proceedings of the National Academy of Sciences, 102(22): 7882-7887.

Winter, A.M., Richter, A., & Eikeset, A.M. (2020). Implications of Allee effects for fisheries management in a changing climate: evidence from Atlantic cod. Ecological Applications, 30(1): e01994.

Wootton, J.T., & Pfister, C.A. (2013). Experimental separation of genetic and demographic factors on extinction risk in wild populations. Ecology, 94(10): 2117-2123.

Xenikoudakis, G., Ersmark, E., Tison, J.L., Waits, L., Kindberg, J., *et al*. (2015). Consequences of a demographic bottleneck on genetic structure and variation in the Scandinavian brown bear. Molecular ecology, 24(13): 3441-3454.

Yamaguchi, K., Nakajima, M., & Taniguchi, N. (2010). Loss of genetic variation and increased population differentiation in geographically peripheral populations of Japanese char *Salvelinus leucomaenis*. Aquaculture, 308: S20-S27.

Young, K. A., Cluney, V. A., & Weir, L. K. (2020). Fisheries‐induced evolution of alternative male life history tactics in Coho salmon. Evolutionary applications, 13(6): 1501-1512.

Zhao, J.L., Wang, W.W., Li, S.F., Cai, W.Q. (2006). Structure of the mitochondrial DNA control region of the sinipercine fishes and their phylogenetic relationship. Acta Genetica Sinica, 33: 793–799.

Declaration of authorship

I hereby declare that I have prepared this thesis independently and without outside assistance. I have not used any sources or aids other than those indicated. The submitted written version of the thesis corresponds to the one on the electronic storage medium.

Furthermore, I certify that this work has not been submitted as a thesis elsewhere.

Date, Signature