Genomic divergence in Atlantic cod populations

Paul R. Berg

Dissertation presented for the degree of

Philosophiae Doctor (PhD)

2016

Centre for Ecological and Evolutionary Synthesis

Department of Biosciences

Faculty of Mathematics and Natural Sciences

University of Oslo

© Paul Ragnar Berg, 2017

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1829

ISSN 1501-7710

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Preface

This dissertation is based on work performed at the Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences at the University of Oslo. Funding for salary was provided through the Department of Biosciences by the Ministry of Education and Research (KD stipend). Research funding was covered by the CodReSeq project, funded by the Research Council of Norway. My supervisors have been Kjetill S. Jakobsen, Sissel Jentoft and Bastiaan Star.
- Kjetill: Thank you for letting me inn under your wings. I am really grateful for all the knowledge you have shared, for great ideas and for fruitful discussions. Last but not least, thank you for really supporting me when things were rough!
- Sissel: Thank you for introducing me to CEES and the cod group and thank you for all the great fishing trips we had. -I will never forget them and I hope there will be more trips in the future!
- Bastian: You are a rock when it comes to scientific thinking and writing. Thank you for all your help on this long and winding road!

Thank you, Nils Christian Stenseth for creating an excellent research environment such as CEES. You have created a dynamic place where the young (and not so young) up-and-coming researchers and students are seen for who they are and encouraged to perform at their very best!

I would also like to thank the entire cod group members (past and present), especially Monica H. Solbakken, Martin Malmstrøm, Alexander J. Nederbragt, Ole K. Tørresen, Helle T. Baalsrud, Michael Matchiner and Julia M. I. Barth. In addition, a special thank you goes to my office mates, Trine B. Rounge who took care of me in the first two year at CEES and Helle T. Baalsrud for bearing with me for the last few years. I would also like to thank the administration and staff at CEES for all help and support.

During my work with this thesis, it was particularly fun to be on sampling trips throughout Norway and I would like to thank everyone participating in these trips and Martin Malmstrøm and Sissel Jentoft in particular (the core Atlantic-cod-sampling-team). However, we could not have done so without the invaluable help from professional fishermen such as Børge Iversen in Ballstad, Lofoten and Darius Ignatavicius at Sørøya Havfiskesenter.

I would also like to extend my gratitude to all national and international collaborators for the good collaboration and access to unique sample material. I wish to thank both co-authors and all those that are acknowledged in the respective papers. Especially I would like to thank Carl André for keeping the Baltic project on track and Christophe Pampoulie, Jeffrey A. Hutchings, Ian R. Bradbury and Halvor Knutsen for fruitful comments, ideas and feedback.

Last but not least, I would like to thank my family and friends. My deepest gratitude goes to my wife Line! Thank you for all your support in these years (and in life in general). Thank you for your understanding when I left for work in the evenings, even though other things needed to be done at home. -I owe you allot! Thank you, mom and dad for putting me into this world and for always supporting me. To all my friends and family: I know that I have been distant and spent all my spare time on this PhD project lately, but all this will change now. I am really looking forward to seeing more of you in the future!

One day, when writing up this thesis, my youngest son said: "Daddy, I really hate your PhD"! I told him that so do I… Even though this sometimes might have been the truth, I must admit that in sum, I really enjoyed riding along this bumpy road. *Max og Julian: Tusen takk for at dere er de dere er. Jeg elsker dere over alt på jord og deres støtte og hengivenhet har mye av æren for at jeg nå endelig er i mål med denne oppgaven. Tusen takk for all kjærlighet og tålmodighetet! –Dere er viktigst av alt*!

List of papers

Paper I

Berg PR, Jentoft S, Star B, Ring KH, Knutsen H, Lien S, Jakobsen KS, André C. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and Evolution* 7 (6): 1644-1663.

Paper II

Berg PR, Star B, Pampoulie C, Sodeland M, Barth JMI, Knutsen H, Jakobsen KS Jentoft S. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific Reports* 6:23246.

Paper III

Sodeland M, Jorde PE, Lien S, Jentoft S, Berg PR, Grove H, Kent MP, Arnyasi M, Olsen EM, Knutsen H. (2016). 'Islands of divergence' in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biology and Evolution* 8 (4): 1012-1022.

Paper IV

Berg PR, Star B, Pampoulie C, Bradbury IR, Bentzen P, Hutchings JA, Jentoft S, Jakobsen KS. (2016). Inversions play a key role in ecotype divergence of Atlantic cod across the Atlantic Ocean. (Submitted manuscript).

Abstract

The core of this thesis has been to address the genomic basis that underlies adaptation to environmental differences in the marine environment. We use Atlantic cod (*Gadus morhua* L.) to enhance our understanding of several ecological and evolutionary questions, where we seek not only to identify the sheer genetic differences between populations, but also seek to identify what these genomic elements are and try to unravel some of the genetic mechanisms that are involved in maintaining and creating such differences.

To put the work into context, I first explain the general principles of population genetics/genomics for non-model species. Furthermore, I describe some key biological features and previous population genetic work in Atlantic cod, which makes a relevant background for interpreting and discussing our results in a general context. The development of a SNParray and linkage maps has provided a valuable tool that is used throughout the thesis. By using this resource in different populations of Atlantic cod, we describe genomic regions likely to be involved in adaptation to different salinity and temperature conditions or differences associated with behavior. The genomic basis of migratory and non-migratory ecotype divergence is explored between the adjacent North East arctic cod and Norwegian coastal cod populations as well as at a trans-Atlantic scale, while adaptation to oceanic and coastal behavioral types is explored at a local scale in the North Sea and at the Skagerrak coast.

Our analyses of these data indicate that a range of genomic regions of several megabases each – which combined span more than 6% of the Atlantic cod genome – play a central role in the genetic divergence between various

populations. Novel findings in this thesis are that some of the most prominent genomic 'islands of divergence' are chromosomal rearrangements in the form of large inversions. Our data suggests a central role for these inversions, each containing hundreds of genes, in maintaining and creating genomic divergence in Atlantic cod. These 'islands' are likely to foster the evolution of co-adapted genes or 'supergenes' by protecting adaptive loci from recombination and thereby facilitating adaptive genomic divergence across different environments and behavioral ecotypes throughout the distribution range. As such, we have provided new insight into the genomic architecture of distinct ecotypes that constitute different life history strategies as well as for populations facing distinctly different environmental conditions. Such knowledge could eventually warrant a more sustainable exploitation and management of Atlantic cod as a species and contribute to a better protection of marine biodiversity in general.

Table of Contents

INTRODUCTION

The fields of ecological and evolutionary genomics seek to understand the genetic mechanisms underlying adaption of organisms to their natural environments. In the marine environment, the potential for such adaptation has historically been thought to be rather low. Many marine species are highly mobile and have pelagic eggs and larvae, characteristics that typically lead to high levels of gene flow that potentially prevent adaptation to local conditions. This view is supported by the typically low observed genetic differentiation between populations of species with such high dispersal capability. Nonetheless, more recently, several studies have reported local adaptation despite high levels of gene flow. These observations are important for management. If hidden population structure is ignored, extinction of locally adapted populations may lead to the irreversible erosion of genomic variation.

Here, we investigate the potential for local adaptation in Atlantic cod (*Gadus morhua* L.), a marine resource of great cultural and economic importance. This is an attractive study system for the following reasons: First it is one of the most studied marine species in the world and its extent of population structure is highly debated (e.g. Bentzen *et al.* 1996; Ruzzante, Taggart & Cook 1998; Knutsen *et al.* 2003; Árnason 2004; Nielsen, Hansen & Meldrup 2006; Nielsen *et al.* 2009c). Second, it has rapidly developing genomic resources such as an annotated genome (Star *et al.* 2011; Tørresen *et al.* 2016). Third, it has pelagic eggs and larval stages, indicating a high gene flow potential. Fourth, it has large population sizes, suggesting that genetic drift is minimal, potentially leading to relatively large amounts of standing genomic variation, which could favor the

potential for local adaptation. Finally, Atlantic cod is a commercially important species under intense harvesting pressure. There is an urgent need to better understand its population structure and connectivity in the marine environment from a basic biological and evolutionary perspective.

Historically, the field of population genetics has focused on the dynamics of only a small number of genetic loci due to methodological limitations. However, genes are not necessarily acting as single units, but could be part of complex genomic structures that are potentially affected and shaped by both evolutionary processes and spatial location within the genome. Knowledge of genes and mutations underlying key adaptive traits is important, but the identification of such genomic regions remains a challenge, particularly in a non-model species such as Atlantic cod. The rapid development of high-throughput sequencing technologies has accelerated genomic research tremendously in the last few years. This makes the collection of individual-based whole-genome data financially feasible, also in non-model organisms (Wheat 2010), and it is now possible to look at traditional population genetic measurements on a continuous scale rather than as point estimates. Specifically in Atlantic cod, the completion of a well-annotated genome assembly has provided us with a valuable resource (Star *et al.* 2011). This has facilitated the identification of genes and genetic variation underlying important traits of ecological and evolutionary relevance, allowing a better separation of the effects invoked by neutral processes from those due to selection.

Population genetics and genomics

Population genetic theory

Population genetic theory postulates that four processes contribute to the genetic variation within populations. These processes are mutation, genetic drift, natural selection and gene flow, each of which is further influenced by effective population size (N_e), population demography, population subdivision, and general life history of the organism (Hartl & Clark 2007). Out of these four evolutionary processes, mutations are often (for simplicity reasons) assumed to not substantially alter the frequencies of existing alleles in large populations (Kimura & Crow 1964; Kimura 1969). Hence, changes in allele frequencies are usually interpreted as the result of genetic drift, gene flow and/or selection within populations. The descent of a given locus in the genome is described by Mendel´s first law (Mendel 1866) and it is often assumed that if individuals have identical alleles, these alleles have a common mutation origin and are identical by descent (IBD) (Wright 1921). Differences in allele and genotype frequencies between populations are further used to infer population structure, gene flow and selection.

Given that both neutral (i.e. genetic drift) and selective evolutionary processes shape the genetic makeup among populations, it is important to disentangle these effects. Genetic divergence at neutral loci between isolated populations is caused by random genetic drift or demographic processes such as population bottlenecks and is expected to affect loci on a genome wide scale. Genetic divergence among populations can be estimated by calculating e.g. F_{ST} values (Wright 1922), which measures population differentiation due to genetic

structure, based on the variance of allele frequencies between populations. Neutral genetic divergence among subpopulations will increase over time and this divergence is inversely related to effective population size (as the frequency of observed heterozygosity is expected to decrease by a factor of $1\text{-}1/(2N_e)$ per generation) in a finite diploid population. Hence, neutral divergence in populations with large effective population sizes can potentially require a substantial timeframe to become established (Hartl & Clark 2007). Over time, the heterozygosity of a locus will decrease towards zero and the F_{ST} value will approach towards one (at neutral loci in two isolated finite populations, due to genetic drift). F_{ST} values that are significantly higher than zero may indicate ecologically relevant population structure. Thus, applying molecular methods to identify such population structure is of special value in commercially important fish species such as Atlantic cod, guiding policy makers in making scientifically sound decisions in fishery management (e.g. Nielsen *et al.* 2001). Such decisions are often made on the background of neutral genetic markers, as genetic markers under selection may reflect environmental differences and not the breeding structure, necessitating analyses of neutrality on the genetic markers used (e.g. Moritz 1994; Laikre, Palm & Ryman 2005). However, some recent research also advocate the use of non-neutral markers in fisheries management and conservation (Nielsen *et al.* 2012; Funk *et al.* 2012).

When natural selection is at work, specific alleles may be favored in different environments. Due to this, different allele frequencies at a locus under selection (or a genomic region under selection) can be caused by different environmental pressure and may not necessarily reflect the population origin (Guinand, Lemaire & Bonhomme 2004) (or the breeding structure) as it would

under neutral genetic conditions. Processes like genetic drift are expected to act equally throughout the genome while directional selection is expected to leave local footprints in specific (and narrower) regions of the genome where the size of the selected region is dependent on the local linkage disequilibrium (see below). As selection in some cases may be a more swift process than random genetic drift, genetic markers under selection may be utilized to examining recent population divergence (Reiss *et al.* 2009). In addition, causes and consequences of selection can be examined to understand biological questions such as local adaptation, adaptation to extrinsic (biotic and abiotic) factors, behavioral traits and factors such as growth rate, age at maturation etc. giving valuable insight into the biology of the species under investigation (e.g. Luikart *et al.* 2003).

Selection and methods for outlier detection

To understand the contribution of molecular adaptation and selection in shaping genomic variation is important, not only to study the genetic basis of adaptation, but also to remove non-neutral loci (i.e. outliers) before computing population genetic parameters that assume neutrality. A growing number of statistical tests and methods, that build upon the increased focus and amount of genomic data in non-model species, are available to detect selection within a species, each with its strengths and weaknesses (e.g. Narum & Hess 2011). As such, several different outlier methods (detecting loci putatively under selection) have extensively been used within population genetics. There are several challenges in outlier detection analyses, including detection of false positives, false negatives and complications due to underlying population structuring. For this

reason, the use of several independent (and often complementary) methods are often used to identify candidate loci under selection and Roesti *et al.* (2012) suggest that uninformative markers at low frequency should be excluded from outlier analyses to increase the power of genome scans. Outlier tests may also detect genomic patterns that are not necessarily caused by selection (or at least selection caused by local adaptation) as the observed patterns could result from endogenous incompatibilities such as underdominance, epistasis or pre- and post-zygotic isolation (Bierne *et al.*, 2011).

One of the first and most commonly used outlier detection approaches was developed by Beaumont and Nichols (1996), which is now implemented in the software LOSITAN (Antao *et al.* 2008). As F_{ST} values are often strongly correlated with heterozygosity at a locus, this can be used to infer the selection status of a locus. Comparisons are made of F_{ST} values in relation to heterozygosity of individual loci, based on a neutral distribution and generated by means of coalescence simulations in a symmetric island migration model at mutation-drift equilibrium. This method is known to be robust within a wide range of non-equilibrium conditions, but it can be sensitive to demographic variations among populations as well as hierarchical genetic structure, that may result in the detection of false outliers (Storz 2005). A more recent outlier detection approach, BAYESCAN, (Foll & Gaggiotti 2008) is a Bayesian regression approach which, based on F_{ST} coefficients, measures the discordance between global and population-specific allele frequencies. The degree of differentiation, based on F_{ST}, is decomposed into a locus-specific component (α), shared by all populations, and a population-specific component (β), shared by all loci. Selection is assumed when alpha is necessary to explain the observed pattern of

diversity. The underlying assumptions for BAYESCAN and LOSITAN are different, as LOSITAN assumes an island model, which may be violated if the neutral population structure is highly divergent. In general, BAYESCAN is assumed to be the more conservative of the two tests described (Narum & Hess 2011).

It has been suggested that outlier tests may produce high false positive rates due to the effects of population demography and bottlenecks (see e.g. Narum & Hess 2011; de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014). One way to reduce this effect is to perform outlier analyses between pairs of populations, since this partly omits the methodological weakness of population structure/demographic processes in the datasets (Vitalis, Dawson & Boursot 2001). In addition, overall divergence based on global outlier tests may not detect candidates that are under selection in only some of the populations (Vitalis, Dawson & Boursot 2001). Hence, specific outlier patterns may be more accurately identified to the population (and in some instances the corresponding environment) being responsible for the observed selection pattern. Another way of minimizing the effect of population structuring is to compensate for population structure by calculating empirical *P*-values. This is done by testing all SNP markers against a null distribution created by a set of putatively neutral SNP markers that are determined a priori (Lotterhos & Whitlock 2014).

Recent simulation studies have suggested that correlation-based approaches in many instances outperforms more traditional outlier detection approaches in accurately identifying loci under divergent selection (de Mita *et al.* 2013; de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014). One such approach (again, among a growing number of methods) is BAYENV (Coop *et al.* 2010; Günther & Coop 2013), which not only identifies correlations between

genotype and environmental factors, but also accounts for demographic signals. Hence, it is possible to disentangle the genetic variation that results from selection from demographically derived patterns in a better way, given that measurable/quantifiable environmental factors are available. The rationale behind this method is to estimate a neutral covariance matrix based on a large set of control loci. Secondly, a test for covariance between the environmental variables and the population specific allele frequencies at each genetic marker is performed, using the neutral covariance matrix as a reference (null model) to control for shared population history and gene flow. For each test, a Bayes factor (BF) is calculated based on the ratio of the posterior probabilities between the two models. A high BF indicates support for the alternative model where the environmental variable has a linear effect on the locus of interest. Even though Coop *et al.* (2010) claim that including a relatively small number of selected loci in the null model will have little impact on the parameter estimates for the null model, using putatively neutral SNPs to parameterize the null model in BAYENV (as recommended by Lotterhos & Whitlock (2014)) may improve this approach even further. While these above approaches represent an increasing set of sophisticated statistical tools to detect selection in genomic data, it is important to acknowledge that such outlier and correlation based approaches have a bias towards detecting selection that has a relatively simple genomic architecture (i.e. does not detect the combined effect of several genetic markers) and predominantly detects outlier loci of large effects. As such, these methods do not detect potentially evolutionary important selection processes resulting from epistasis or polygenic loci.

Linkage disequilibrium

The outlier methods described above are treating each locus separately and ignore the effects from multiple loci that are non-randomly associated with each other. An important step in considering the effects from multiple loci acting together, is to calculate and investigate linkage disequilibrium (LD), which is described as non-random association of two or more loci that deviate from statistical equilibrium expectations (e.g. Lewontin & Kojima 1960; Hill & Robertson 1968; Slatkin 2008). Linkage disequilibrium is a key factor in evolutionary biology as genome wide LD can reflect population history, patterns of geographic subdivision and demographic events whereas LD in localized genomic regions reflects the forces that causes allele frequency divergence such as natural selection, gene conversion and mutations, which in turn are dependent on local recombination rates (Slatkin 2008). Mapping and quantifying LD across a genome have importance in disentangling the effects of demography and selection (Cutter 2006) and in addressing genomic causes and consequences of e.g. local adaptation (Ellegren & Sheldon 2008; Slate *et al.* 2009). As such, there is an association between LD and signatures of divergent selection, which is described in the next two sections.

Islands of divergence

In the process of population divergence and speciation, heterogeneous genomic divergence can be formed (Nosil, Funk & Ortiz-Barrientos 2009; Smadja & Butlin 2011), leaving footprints in the form of elevated levels of divergence in genomic regions under selection. Physically linked loci and mutations in close proximity to the loci under selection can show similar divergence via hitchhiking (c.f.

Charlesworth, Nordborg & Charlesworth 1997; Via 2009; 2012). As a result, the size of the local genomic island can increase considerably (Smith & Haigh 1974; Feder *et al.* 2012b; Feder, Egan & Nosil 2012a).

During speciation and population divergence in the presence of gene flow (reviewed in: Tigano & Friesen 2016), the establishment and maintenance of genomic regions that can sufficiently resist gene flow is only likely if divergent selection (or sexual selection) is strong and hence the initial barriers to gene flow are likely to evolve quickly (Via 2001; Hendry, Nosil & Rieseberg 2007). As a result of such strong selection, the genomic regions causing the reproductive isolation become particularly distinctive relative to the remaining genome, facilitating its discovery in empirical analyses (Via 2009). Hence, studying the genetic changes that contribute to reproductive isolation in partly reproductively isolated populations before they become confounded by additional genetic differences towards the end of the speciation process, may reveal important aspects of the divergence process (Via 2009). In such a scenario, which is consistent with the early stages of genetic divergence where a substantial amount of gene flow is still prevalent, one would expect the presence of relatively few but potentially large genomic islands of divergence (Yeaman & Whitlock 2011; Feder *et al.* 2012a). This pattern of early divergence has been shown in studies of e.g. pea aphids (*Acyrthosiphon pisum pisum*), African malaria mosquitoes (*Anopheles gambiae*) and *Heliconius* butterflies (Turner, Hahn & Nuzhdin 2005; Via & West 2008; Via 2009; Nadeau *et al.* 2012; Via *et al.* 2012).

Over time, recombination will reduce LD across the genome, resulting in larger patterns of genome-wide divergence, although heterogeneity among smaller regions may still be present due to varying degree of selection and

recombination. As a result, ecologically favored alleles will predominate in one habitat and neutral- and universally favored alleles will potentially be present in all habitats (e.g. Savolainen, Lascoux & Merilä 2013). At this point, fixed differences can be observed between populations (Feder *et al.* 2012a; Seehausen *et al.* 2014). This is consistent with the theory of isolation by adaptation (Orsini *et al.* 2013) and the later stages of ecological speciation, where gene flow is small or non-existent (Feder *et al.* 2012a). Such patterns of divergence have previously been identified in other fish species like three-spine sticklebacks (*Gasterosteus aculeatus*) (Roesti *et al.* 2012; 2014) and lake whitefish (*Coregonus clupeaformis*) (Gagnaire *et al.* 2013).

Inversions

Genomic islands of divergence (Wu 2001; Nosil *et al.* 2009), as the ones described above, consist of linked loci within genomic regions under selection and are known to emerge through e.g. divergence hitchhiking (Via 2012). However, such patterns can also emerge through other processes that reduce recombination in genomic regions, such as chromosomal rearrangements (Kirkpatrick & Barton 2006) where distinct LD blocks will contain the entire selected region. In such chromosomal rearrangements, the rate of crossing over is reduced by several orders of magnitude (Feder, Nosil & Flaxman 2014), and therefore large genomic regions may be affected (Rieseberg 2001; Noor *et al.* 2001). It has been suggested that chromosomal rearrangements in the form of large inversions can play a vital role in maintaining polymorphism in complex traits and play a key role in the genomic process involved in local adaptation (Yeaman & Whitlock 2011). A suggested mechanism is that an inversion captures

several locally adapted alleles, as the inversion suppresses meiotic recombination (Fig. 1) in heterozygous individuals (Rieseberg 2001; Kirkpatrick & Barton 2006), forming so called 'supergenes' (e.g. Joron *et al.* 2011; Schwander, Libbrecht & Keller 2014; Thompson & Jiggins 2014).

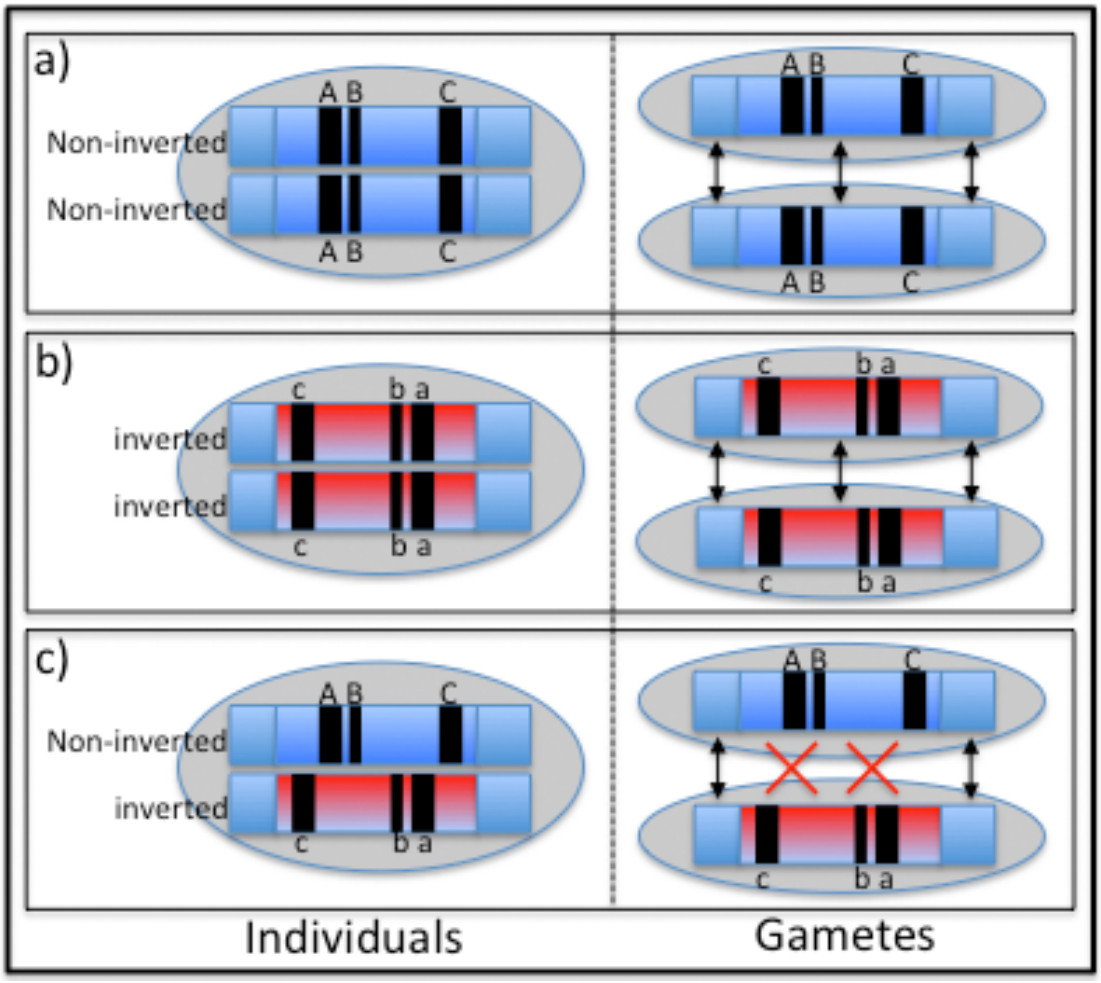

Figure 1. A schematic overview on how inversions may suppress meiotic recombination in heterozygous individuals. A potentially inverted chromosomal segment containing three different loci, segregating for the alleles A/a, B/b and C/c are illustrated. The inverted segment is depicted in red color and the black arrows and red crosses indicate regions where recombination can, and cannot occur respectively. In the homozygous non-inverted (a) and inverted (b) variants, the gametes are free to recombine during meiosis throughout the illustrated area. In the individual that is heterozygote for the inversion (c), the gametes are free to recombine only in the regions before and after the inversion segment. Within the inversion segment, recombination is suppressed between gametes due to inconsistent pairing of homologous regions during the meiosis.

Indeed, in some species strong signatures of selection have been shown to be associated with chromosome inversions (e.g. Jones *et al.* 2012a; Twyford & Friedman 2015; Zinzow-Kramer *et al.* 2015) but few studies have identified the

actual target genes for selection within these inversions (Kirkpatrick & Kern 2012) and the potential for adaptation as well as the underlying genetic architecture remain unclear in most species. Moreover, it is not necessarily the genes (i.e.; the protein coding regions) within an inversion that is under selection. In some cases, it might be the break points of the inversion that are locally adapted, working through disruption of gene expression (caused by the inversion) causing selection to favor locally adapted break points (Puig, Cáceres & Ruiz 2004; Matzkin *et al.* 2005). An increasing number of studies however, support the prominent role of inversions in creating and maintaining locally adapted genomic regions in heterogeneous environments (e.g. Rieseberg 2001; Kirkpatrick & Barton 2006; Schaeffer 2008).

Atlantic cod (*Gadus morhua*)

Biology of Atlantic cod

The Atlantic cod has been, and still is one of the most commercially important fish species in the Atlantic Ocean. In fact, it is one of the most important commercial fish species in human history and it has been harvested on both sides of the Atlantic Ocean since the Viking age (Kurlansky 1997). In the last few decades, heavy exploitation and overfishing has led to severe decline in several important cod populations (Caddy & Cochrane 2001) and to collapses of several major cod fisheries (Hutchings 2000), with almost no recovery despite a strict regulation and suspension of fishing activities in these areas. Within the same timeframe, profound phenotypic changes have been recorded in several commercial cod stocks (Olsen *et al.* 2005; 2009; Harrald, Wright & Neat 2010), which in some instances have been implicated as an effect of fisheries induced evolution (FIE) (Olsen *et al.* 2004; Swain, Sinclair & Hanson 2007). As fisheries usually are size selective and hence harvest primarily large (and potentially fast growing) individuals, combined with the fact that fisheries often represents the most dominant mortality factor in these species (Stenseth & Dunlop 2009), selective harvesting can change the evolutionary trajectory of heavily exploited species. At the same time, ocean climates are changing at unprecedented speed, resulting in unpredictable consequences for both ecosystems and fisheries (Gewin 2015). As such, rapid ocean warming have also been suggested as the cause of some of the population collapses in the Northwest Atlantic (Pershing *et al.* 2015), leaving the cod with too little time to adapt to changing conditions.

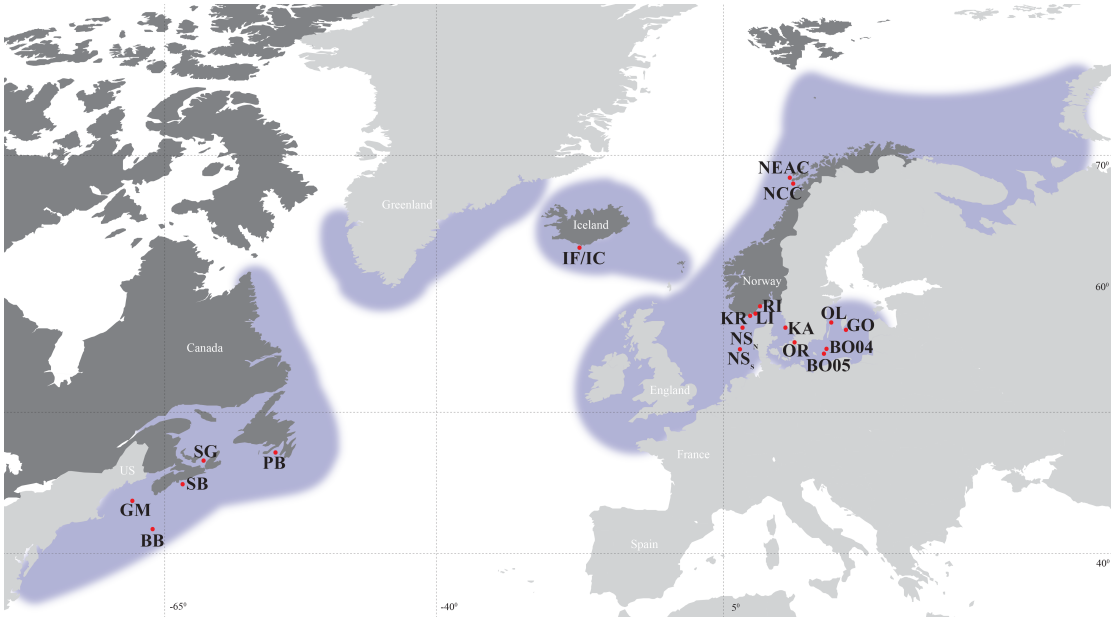

Figure 2. The distribution range of Atlantic cod and the sampling locations of Atlantic cod used in this thesis. Blue shaded areas depict the distribution range of Atlantic cod. Red dots indicate the positions where samples were collected. The Risør, Lillesand and Kristiansand collections consist of an inner and an outer location and the Icelandic samples were collected at several locations in the waters around Iceland. The Icelandic samples were later categorized as Frontal or Coastal, based on Data Storage Tag profiling. See Thorsteinsson *et al.* (2012) and Sodeland *et al.* (2016) for a detailed view of the Icelandic and Skagerrak coast sampling localities, respectively.
PB = Placentia Bay, SG = Southern Gulf of St. Lawrence, SB = Sambro, GM = Gulf of Maine, BB = Browns Bank, IF = Iceland Frontal, IC = Iceland Coastal, NEAC = Northeast Arctic cod, NCC = Norwegian coastal cod, NS_S = North Sea south, NS_N = North Sea north, KR = Kristiansand, LI = Lillesand, RI = Risør, KA = Kattegat, OR = Øresund, BO04 = Bornholm 2004, BO05 = Bornholm 2005, GO = Gotland and OL = Öland.

Atlantic cod inhabit the temperate continental shelves, banks and coastal regions in the North Atlantic Ocean. Its distribution (Fig. 2) ranges from the English channel and New England in the south to Greenland and the Barents Sea in the north (Mieszkowska *et al.* 2009). It is distributed along a variety of habitats, from the shoreline down to deeper waters (>600 meters) off the continental shelf and within fjords and estuaries. Atlantic cod exploits a variety of different environments, such as saline environments in the North Sea and the Barents Sea and low saline environments in the Baltic Sea, the White Sea, in inner fjord locations and within a few landlocked locations (Hardie & Hutchings 2011). In addition, it is experiencing a wide range of temperature differences as e.g. the

North east arctic cod (NEAC) stay at relatively warm winter ground in the Barents Sea and move to colder summer feeding grounds (Bergstad, Jørgensen & Dragesund 1987; Jørgensen 1992; Ottersen, Michalsen & Nakken 1998). The Baltic cod experiences large annual temperature differences, ranging from approximately 2 - 18 °C within a year (Rak & Wieczorek 2012) and large temperature fluctuations are also observed in the southern parts of the distribution range (Righton *et al.* 2010). Atlantic cod grow throughout its lifespan and can reach a length of more than two meters and live approximately 25 years (Salvanes, Skjæraasen & Nilsen 2004). Atlantic cod is a batch spawner that is extremely fecund, has pelagic eggs and larvae and large effective population sizes (Ward, Woodwark & Skibinski 1994). As the cod grows older, a larger proportion of its available energy resources is allocated into reproduction (as opposed to allocated to growth prior to maturation) resulting in an increase of reproduction with age (Marteinsdottir & Begg 2002).

The NEAC population supports the renowned "skrei" fisheries and its behavior is characterized by long distance migrations from the spawning grounds along the Norwegian coast to feeding areas in the Barents Sea. The main spawning grounds are off the Lofoten islands (Brander 2005) and after spawning, the majority of eggs and larvae drift along the coast into the nursery area in the Barents Sea. In contrast, Norwegian coastal cod (NCC) inhabits coastal- and fjord areas along the Norwegian coast, perform relatively short coastal migrations (Hylen 1964; but see Michalsen *et al.* 2014), and spawn along most of the Norwegian coast (Jakobsen 1987), including the Lofoten area (Nordeide 1998). In Iceland, the main spawning grounds for the cod populations are of the Southwestern coast and the eggs and larvae drift clockwise around

Iceland towards nursery areas of the Northern coast (Jónsson & Valdimarsson 2005) but spawning is also known from fjord locations around the island (Begg & Marteinsdottir 2000). In some years, eggs and larvae may drift to the Greenland East coast, depending on the strength of the North Atlantic current (Jónsson & Valdimarsson 2005). In the Northwest Atlantic, distinct spawning stocks are known all along the coast and also at offshore locations such as e.g. Flemish Cap (Ruzzante *et al.* 1998; 2001; Beacham *et al.* 2002). The cod stocks from the Northern Gulf of St. Lawrence are known to undertake extensive annual migrations, spending winter at deep waters off the southern coast of Newfoundland and moving towards the spawning grounds on the west coast of Newfoundland in late spring (Brander 2005). Also the cod stocks from Southern Gulf of St. Lawrence are considered to be discrete from adjoining areas and are known to perform annual feeding migration (Brander 2005). In the south, distinct populations are known from Gulf of Maine, George Bank and Southern New England and these are known to be mainly resident or display only low level of movement (Howell *et al.* 2008).

Population genetic studies of Atlantic cod

Historically, the potential for marine species to adapt to their local environment, have received little attention. From the 1960s when population genetic analyses in marine species emerged, a paradigm persisted for decades, claiming that most marine fish species consisted of large undifferentiated populations. In the following years, most studies described the homogenizing effects of gene flow in the absence of obvious physical barriers in marine environments, showing low degree of genetic differentiation among populations and high degree of within-

population variation (Gyllensten 1985; Ward *et al.* 1994). Research focused primarily on genetic variation in presumably neutral genetic markers and the effects of selection were not given much attention (Hauser & Carvalho 2008).

Already in the 1930ies, two distinct groups of Atlantic cod along the Norwegian coast were described based on otolith morphology and growth zones by Rollefsen (1933). Since then, attempts to describe population structuring in Atlantic cod have been a focus for several decades and today well over 100 studies, using population genetic methodology, have been published for this species. Despite the fact that no other commercial fish species have been investigated to this degree within a population genetic framework (Reiss *et al.* 2009), the degree of population structuring (Reiss *et al.* 2009) and adaptive structuring (Hutchings *et al.* 2007) still remains debated. In Atlantic cod, several studies suggest weak spatial structuring over relatively large areas e.g. (O'Leary *et al.* 2007; Reiss *et al.* 2009) while other studies have detected loci that seems to have a clear clinal structure and/or elevated divergence e.g. (Pogson 2001; Case *et al.* 2005; Nielsen *et al.* 2006), supporting adaptation and spatial structuring in Atlantic cod populations.

Some of the first population genetic studies on Atlantic cod used hemoglobin polymorphisms to identify population structuring (Sick 1961; 1965) and a clinal variation in hemoglobin and blood group (serum transferrin) variation were detected along the Norwegian coast (Frydenberg *et al.* 1965; Moller 1968) and the observed patterns were explained as population subdivision in several papers (e.g. Sick 1961; 1965; Frydenberg *et al.* 1965; Moller 1968). This observed spatial variation could however be the results of natural selection in a heterogeneous environment and not neutral population

structuring, as pointed out in a book by Williams (1975). Karpov & Nivkov (1980) showed different oxygen carrying capacity of the different hemoglobin forms in cod populations that were residing in areas of different temperature regimes. Later, it has been well documented that different hemoglobin variants are preferred in cold and more temperate waters (Brix, Thorkildsen & Colosimo 2004; Imsland *et al.* 2004; Andersen *et al.* 2009; Borza *et al.* 2009; Andersen 2012). In 1985, Mork *et al*. (1985) did a study on Atlantic cod, using allozymes, and concluded that there were little genetic differentiation throughout the distribution range except for relatively large genetic distances between Atlantic and Baltic cod. Later on, the use of mitochondrial DNA (mtDNA) have revealed several instances of fine scale genetic structure, but also high levels of gene flow (e.g. Carr & Marshall 1991; Árnason, Palsson & Arason 1992; Carr *et al.* 1995; Árnason & Palsson 1996; Árnason, Petersen & Palsson 1998) in addition to a broad trans-Atlantic cline with shallow gene genealogy within the two regions (Árnason 2004).

An important point in the history of population genetic analyses of Atlantic cod was the discovery and use of polymorphic sites within or close to the pantophysin gene (*Pan* I/*Syp* I) (Fevolden & Pogson 1996; 1997). This gene has later been used in a wide range of population genetic studies (e.g. Karlsson & Mork 2003; Sarvas & Fevolden 2005; Pampoulie *et al.* 2006; Jakobsdóttir *et al.* 2011; Fevolden *et al.* 2012; Makeenko *et al.* 2014; Andersen *et al.* 2015) and is also commonly used to determine individuals as NEAC or NCC (Fevolden & Pogson 1997). However, more recent studies indicated that genetic variation at the *Pan* I locus may be affected by natural selection (Karlsson & Mork 2003; Pogson & Mesa 2004; Árnason, Hernandez & Kristinsson 2009) even though the

function of the gene is largely unknown (but see Andersen *et al.* 2015).

Population genomics and adaptive genomic variation

In the last decade, single nucleotide polymorphisms (SNPs) have replaced mtDNA and microsatellites as the preferred choice of genetic markers to study population structuring and genomic divergence in many wild species or populations. Consequently, a growing number of SNP based studies have been performed in Atlantic cod, using a moderate number of markers (up to a few hundred SNPs) (Moen *et al.* 2007; Nielsen *et al.* 2009b; Bradbury *et al.* 2010; 2013; Hemmer-Hansen *et al.* 2013). Further, a shift in focus has emerged towards an increased attention on identification and description of adaptive variation and to detect genomic regions underlying local adaptation (e.g. Cano *et al.* 2008; Hauser & Carvalho 2008; Nielsen *et al.* 2009a), rather than on neutral genetic variance. Several of these studies focused on only one or a limited number of candidate genes or divergent loci (Hemmer-Hansen *et al.* 2007; Bradbury *et al.* 2010; André *et al.* 2011), hence, genome-wide patterns of divergence cannot be explored thoroughly. However, there is a steadily growing body of articles focusing on genome wide genomic divergence in aquatic organisms (e.g. Feulner *et al.* 2013; Karlsen *et al.* 2013; Tine *et al.* 2014; Pujolar *et al.* 2014; Jacobsen *et al.* 2014; Guo *et al.* 2014; Guo, Li & Merilä 2016). These studies not only include a large number of genetic markers, they also include large sets of candidate genes and identify outlier loci that can enlighten our perception of how genetic differentiation is distributed throughout the genome and how the genome is affected by the opposing forces of selection and gene flow. A general drawback on most studies conducted at a genome-wide scale is

that only a limited number of populations have been investigated which impedes a proper inference on the association between genotype and environment.

With the completion of a fully sequenced and annotated Atlantic cod genome (Star *et al.* 2011), combined with the rapid development (and reduced costs) of next-generation sequencing technology, a whole new era has opened up for population genomic work in Atlantic cod. In addition, important knowledge on the unique immune system of Atlantic cod and other gadoids have recently been thoroughly examined (Solbakken *et al.* 2016a; b), based on the completion of the Atlantic cod genome and a range of other *de novo* sequenced gadoid genomes. Even though it is in its infancy, we are now able to use this invaluable resource, and also a new and significantly improved cod genome (Tørresen *et al.* 2016), to face many of the questions and challenges addressed in the earlier population genetic studies on Atlantic cod.

SNPchip and linkage maps

To increase the SNP coverage, and hence get a better resolution to our analyses, we constructed a custom Illumina SNP-array that could be used for multiple purposes. As such, SNPs were selected so that; (1) they could be used for genetic map construction; (2) they could be used as markers for different genomic regions and; (3) they could be used for general population genetic purposes in Atlantic cod. The SNP-chip was designed and constructed as part of the Norwegian Cod SNP Consortium (CSC), which consists of representatives from four Norwegian research organizations (University of Oslo, Norwegian University of Life Sciences, Institute of Marine Research, and NOFIMA). To do so, genomes from seven Atlantic cod individuals, collected from a wide geographic

range across the Northeast Atlantic (from the following locations: Risør, North Sea/ English Channel, the Baltic Sea, the Varangerfjord, the Ullsfjord, the Viking bank and Lillesand) were shotgun sequenced as paired-end reads, using Illumina GAii instrumentation. For each sample an average of 79% of reads were aligned to the reference genome (Star *et al.* 2011) using the Burrows–Wheeler Aligner (Li & Durbin 2009). From this, SNPs were detected using SAMtools (Li *et al.* 2009) and the list of 2,877,794 putative SNPs was reduced using a variety of filters such as physical distribution, functional associations, and minor-allele frequency > 0.1 in the sequenced samples. From a total of 10,605 on the final SNP-Chip, genotypes were checked for Mendelian errors in a large family material and quality and validity assessment were performed specifically for the different projects. All SNPs were submitted to dbSNP (www.ncbi.nlm.nih.gov/ SNP) and are available by their ss# or rs# accession numbers. Genotypes were clustered using the Illumina GenomeStudio software 2011.1, and only SNPs found to be polymorphic loci, inherited in a Mendelian way and showed good clustering in >95% of the individuals in the respective projects were included for further analyses (see the respective papers for details). A publication on the development of this 12K SNP-Chip in Atlantic cod is currently under preparation (Kent M, Kirubakaran TG, Berg PR, Baranski M, Dahle K, Jakobsen KS, Jentoft S, Johansen T, Nederbragt AJ, Nome T, Star B & Lien S).

A linkage map, which is a genetic map showing the position of genetic markers along each linkage group relative to each other (in terms of recombination frequencies and not physical distance), based on 924 SNPs has been available for quite some time in Atlantic cod (Hubert *et al.* 2010). However, with increased SNP coverage and the advent of new sequencing technology, a

denser and updated linkage map is warranted. Hence, detailed linkage maps for Atlantic cod, based on the 12K SNP-Chip is currently under preparation (Grove H, Kirubakaran TG, Kent M, Baranski M, Nome T, Sandve S, Berg PR, Sodeland M, Dahle G, Sonesson A, Johansen T, Andersen Ø & Lien S) and a preliminary version of this linkage map was used in the papers of this thesis.

OBJECTIVES

Our primary goal was to address genome-wide genetic variation and differentiation in Atlantic cod that were subject to several different environments and habitats. By taking advantage of the information from the annotated Atlantic cod genome assembly (Star *et al.* 2011) and by utilizing a wide range of Atlantic cod populations, we seek to address general ecological and evolutionary topics such as identifying the genomic context that differs between populations and ecotypes, and to identify whether such differences may be important in an evolutionary context. Of particular interest was to identify genomic regions displaying footprints of selection associated with different environmental regimes and migratory/non-migratory behavior and to compare these patterns between different sets of populations. Further, we aimed at investigating whether divergence between distinct populations and ecotypes were mainly caused by natural selection or by neutral processes and whether exposure to similar environmental factors lead to the evolution of similar genomic adaptation patterns, hence shaping the genomic architecture of Atlantic cod. The objectives for the separate papers are described in the next section.

SHORT DESCRIPTION OF PAPER I - IV

Paper I

Berg PR, Jentoft S, Star B, Ring KH, Knutsen H, Lien S, Jakobsen KS, André C. (2015). Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and Evolution* 7 (6): 1644-1663.

We here provide the first individually genotyped genome-wide approach to date on this species. We investigated the genomic signatures of local adaptation in Atlantic cod, along a natural salinity gradient, ranging from 35‰ in the North Sea to 7‰ within the Baltic Sea. As the Baltic Sea originated approximately 8,000 years ago, adaptation to low saline conditions is likely to be of relatively recent evolutionary origin, providing an excellent opportunity to study the genomic architecture behind salinity adaptation in a natural environment. Our main objective was to identify directionally selected loci that were correlated with habitat differences in salinity, oxygen and temperature. In addition, we sought to identify candidate genes for adaptation to low salinity conditions, i.e. genes affecting the individuals both at the egg and larval stage as well as at the adult stage. Overall, our results show that regions, consisting of directionally selected loci, are strongly correlated with habitat differences in salinity, oxygen and temperature and several outlier SNPs reside within ecologically important genes that could affect egg buoyancy and general osmoregulation. As ecological adaptation to a low-saline environment may contribute to reduced gene flow and thereby promote population divergence, investigation of the genomic architecture of Baltic cod supply insights into ecological speciation in nature, and especially the genetic link between adaptation and reproductive isolation.

Paper II

Berg PR, Star B, Pampoulie C, Sodeland M, Barth JMI, Knutsen H, Jakobsen KS Jentoft S. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific Reports* 6: 23246.

Here we address one of the most longstanding controversies in Atlantic cod research, namely the population structure of - and the connectivity between - Northeast Arctic cod (NEAC) and Norwegian coastal cod (NCC). The genomic architecture underlying behavioral ecotypes (migratory/non-migratory) have so far been unclear. However, distinct phenotypic differences between migratory and non-migratory ecotypes, which spawn in the vicinity of each other, offers an excellent opportunity to explore such patterns within a framework where the potential for both natural selection and gene flow is high. As such, our main objective was to investigate the genomic basis for migratory and non-migratory behavior. Based on the fact that regions of elevated genomic divergence has previously been detected in other Atlantic cod populations, an additional aim was to show that these regions are most likely genomic rearrangements in the form of large inversions. The frequencies of these distinct regions differ markedly between migratory and non-migratory ecotypes and the observed patterns strongly suggest that these chromosomal rearrangements are instrumental in local adaptation and separation of Atlantic cod populations, leaving footprints of large genomic regions under selection. Our findings demonstrate the power of using genomic information in further understanding the population dynamics and defining management units in one of the world's most economically important marine resources.

Paper III

Sodeland M, Jorde PE, Lien S, Jentoft S, Berg PR, Grove H, Kent MP, Arnyasi M, Olsen EM, Knutsen H. (2016). 'Islands of divergence' in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biology and Evolution* 8 (4): 1012-1022.

Here we expanded our previous knowledge about the importance of chromosomal rearrangements for adaptation. Our main objectives were to investigate the genomic basis of adaptation to oceanic and coastal behavioral types in the North Sea/Skagerrak area. We do this by addressing patterns of genomic differentiation on a fine geographic scale in three parallel systems, which provides a good system to distinguish processes that act similarly on all populations from processes acting specific on populations in a distinct environment. We observe patterns of recombination and divergence that resembles genomic signatures described previously for large polymorphic inversions. Intriguingly, the pattern of spatial genetic structure indicated that all of the inner coastal samples were similar to each other (even though they were located in distinctly different fjord systems), but divergent from the outer coastal samples, which again were similar to each other. Differences in rearrangement frequencies between coastal and oceanic environments were observed in two linkage groups and outer coastal samples were found intermediate to the oceanic and the inner coastal samples, Combined, these results suggest a role for these rearrangements in ecological adaptation.

Paper IV

Berg PR, Star B, Pampoulie C, Bradbury IR, Bentzen P, Hutchings JA, Jentoft S, Jakobsen KS. (2016). Inversions play a key role in ecotype divergence of Atlantic cod across the Atlantic Ocean. (Submitted manuscript).

In this paper, we extend on the results from paper II and sought to address genomic differentiation of Atlantic cod at a trans-Atlantic scale. Our main objective was to investigate whether large chromosomal rearrangements that are known do differentiate between migratory and non-migratory ecotypes in Northeast Atlantic could be a distribution-wide phenomenon. Specifically, we investigate whether the 'islands of divergence' that has previously been observed in Northwest Atlantic cod consists of inversions such as the ones observed in Northeast Atlantic cod. Overall, our data suggest a central role for several large chromosomal inversions, each containing hundreds of genes, in sustaining and creating genomic divergence in Atlantic cod on both sides of the Atlantic Ocean. We hypothesize that the same inversions contribute to genetic divergence among Northwest- and Northeast Atlantic cod, which separated more than 100,000 years ago, supporting a common origin of these inversions, predating the trans-Atlantic split. By doing so, we also sought to get new insights into the age of these rearrangements. Moreover, it appears that the non-migratory ecotype is always dominated by the ancestral collinear inversion genotype, containing the highest nucleotide diversity, and that distinct migratory behaviour seems to be derived from non-migratory ecotypes in Atlantic cod. The results have implications for management of Atlantic cod as a species but also provide insight into the process of genomic divergence in marine fish species in general.

DISCUSSION

In times where extinction rates are magnitudes higher than the normal background rate (De Vos *et al.* 2015), understanding of biological processes and diversity is highly needed for conservation and development of sustainable management programs. In the marine environment, species and populations are threatened not only by intense commercial harvesting, but also by global warming, acidification, oil spill, plastic and general pollution. As a consequence, the need to manage the marine environment in a sustainable way is increasing, warranting enhanced understanding of population connectivity and structuring from a basic biological and evolutionary perspective. Here, we use Atlantic cod to address general ecological and evolutionary questions such as; what are the genomic differences between populations, and what are the underlying mechanisms responsible for establishing and maintaining these differences? Such questions are intimately connected to intraspecific variation, which we analyze in our respective datasets. As such, the knowledge we generate is of fundamental importance to our understanding of connectivity and adaptation in the marine environment in general, and in Atlantic cod in particular. Increased knowledge of Atlantic cod biology also has a commercial potential, as knowledge about specific cod populations and how they are connected could directly benefit the fishing industry in order to harvest this economically important resource in a more sustainable way. Additionally, insight based on genomic data could be of a commercial interest if genes involved in e.g. growth and maturation are identified, as the use of such information could be beneficial in potential aquaculture initiatives within this species.

At the time when paper I was written, the presence of inversions in the Atlantic cod genome had not yet been documented. However, large 'islands of divergence' were already known to be present in the Atlantic cod genome (Bradbury *et al.* 2013; Hemmer-Hansen *et al.* 2013). As such, our analyses in paper I revolve around 'islands of divergence' in a divergence hitchhiking scenario, although the possibility that inversions are involved in the process were briefly discussed. In paper II, it became evident that chromosomal rearrangements were involved in the genomic differentiation between migratory (NEAC) and non-migratory (NCC) populations of Atlantic cod. Following this, we investigated the genomic divergence in paper III and IV in the light of chromosomal rearrangements. Even though chromosomal rearrangements now have been identified in several Atlantic cod populations (paper II, III and IV), this does not change the conclusions in paper I, even if the origin of the observed 'islands' could have a different explanation than in a divergence hitchhiking scenario. Natural selection is still likely to be responsible for the observed patterns and hence shape population structure on short spatial scales, despite (or possibly because of) the high dispersal capacity of marine organisms, causing a high potential for gene flow (see e.g. Gaggiotti *et al.* 2009; Bradbury *et al.* 2010; André *et al.* 2011; Lamichhaney *et al.* 2012; Bradbury *et al.* 2013; Hemmer-Hansen *et al.* 2013; Defaveri *et al.* 2013).

In paper I, we describe candidate genes both within and outside of the large 'islands of divergence' in great detail. These genes are described as potential targets of selection, likely to be associated with the examined environmental variables such as salinity, temperature and oxygen level and the identified genes may have a relevant function relative to the environmental

differences that the populations experiences. In general, 'islands of divergence' may be the product of processes that act similarly on all populations or are specific to populations in a certain environment (hence local adaptation). To unravel such effects from each other, studies of parallel systems has been used, particularly so in sticklebacks where multiple independent freshwater colonization events from marine founders has been observed in several instances (e.g. Hohenlohe *et al.* 2010; Jones *et al.* 2012a; Deagle *et al.* 2013; Terekhanova *et al.* 2014), indicating parallel evolution. In other instances, such parallelism has been detected to a lesser degree (Defaveri *et al.* 2011; 2013), which possibly indicates the presence of a more regional effect. To do such parallel studies in Atlantic cod may prove difficult for the salinity adaptation described in paper I, but comparisons between the low saline White Sea and adjacent offshore locations could be an exciting parallel study, as could comparisons between oceanic populations and low saline fjord populations (Barth *et al.* in prep). Additionally, arctic lake populations of Atlantic cod (Hardie, Gillett & Hutchings 2006; Hardie & Hutchings 2011) which experience low saline conditions, could be investigated, but such analyses are likely to be hampered by low genetic diversity within these lakes due to potential inbreeding and small population sizes. Parallel studies addressing factors such as temperature, spawning time, migration patterns etc. could be performed by carefully selecting relevant populations with the desired traits resulting from different local adaptations in the same way as we do in paper IV, where similar behavioral traits are investigated on both sides of the Atlantic Ocean.

The continuously growing number of scientific papers describing inversions and the effects that inversion polymorphism has on genomic

diversification and speciation, indicate that inversions and inversion polymorphisms are common phenomena. Due to a growing number of readily available annotated genomes, also the genetic content of such inversions can now be characterized in a straightforward way, and we do that in paper II and IV. However, identifying the actual targets of selection within inversions or other tightly linked genomic regions, is mostly yet to be determined (Kirkpatrick & Kern 2012), also in Atlantic cod. As a consequence, the ultimate mechanisms underpinning local adaptation associated with inversions are often unknown. As an example, in paper II (Berg *et al.* (2016)), we showed that *Pan* I is only one out of approximately 785 genes within a large linked outlier region in linkage group (LG) 1 in the Atlantic cod genome. Since this divergent region in LG1 appears to be inherited as one large rearranged region, it is not necessarily the *Pan* I locus that is under selection as any of the other genes (or a combination of genes) within this region could be the actual target of selection. Other genes within this genomic region have also been examined as likely targets of selection, such as rhodopsin (Pampoulie *et al.* 2015). Hence, further research is needed to unravel the actual targets of selection within this and other such highly linked outlier regions. The difficulties in identifying the actual targets of selection within linked regions are caused by the very nature of the inversions, namely that recombination is reduced within polymorphic chromosomal inversions (see Introduction; Fig. 1), causing difficulties in distinguishing the true targets of positive selection from linked false positive signals. As such, traditional F_{ST} based outlier tests such as FDIST (Beaumont & Nichols 1996)/Lositan (Antao *et al.* 2008), BAYESCAN (Foll & Gaggiotti 2008) and FLK (Bonhomme *et al.* 2010) perform poorly within inversions, due to high baseline divergence between

inverted and non-inverted genomic regions (see Cheng *et al.* 2012). In addition, the detectability of selected loci within an inversion is dependent on the neutral divergence within the inversion, i.e. the timeframe for neutral homogenization via gene flux; see (Guerrero, Rousset & Kirkpatrick 2012) hence violating the underlying assumptions of the standard neutral parameterization methods described above. Assuming that phenotypic differences among ecotypes are present at gene expression level, the same limitations may not be as severe when comparing transcription profiles to identify candidate genes (Mack, Campbell & Nachman 2016; Fuller *et al.* 2016), which then offers an alternative approach to find causative genes and polymorphisms. Validation of candidate genes can also be performed using QTL mapping from natural pedigrees or genetic crosses (Laporte *et al.* 2015), where co-localization of relevant QTLs and genomic islands strengthens the conclusions drawn from the genetic composition underlying 'islands of divergence'. Enrichment tests, searching for GO (gene ontology) terms that are over-represented within inversions relative to the remaining parts of the genome could also be an alternative approach, however when the inversions are large and the GO landscape is not sufficiently well described, as is the case for Atlantic cod (one would need to go through stickleback, zebra fish or even human GO terms), this may not yet be very fruitful. A detected association between a phenotype or an environmental factor and loci under selection does not prove a causal effect of the underlying genes or the genomic region and are hence only indications of causality (Wolf & Ellegren 2016).

Ever since seasonal changes in polymorphic inversion frequencies were observed in *Drosophila* (Dobzhansky 1943) the effects of reduced recombination rates within inversions have been linked to adaptation with gene flow, supported

by subsequent research on e.g. *Drosophila* (Noor *et al.* 2001), *Helianthus* sunflowers (Rieseberg 2001), *Anopheles* mosquitoes (Ayala & Coluzzi 2005) and *Agrodiaetus* butterflies (Kandul, Lukhtanov & Pierce 2007). These papers all show that sympatric species exhibit more differences associated with inversions than allopatric species. Moreover, recent research on e.g. tropical reef fishes (Martinez *et al.* 2015) and a finch species complex (Hooper & Price 2015) indicate a quicker fixation of inversions in lineages with higher dispersal potential and gene flow. This is consistent with theory predicting that gene flow favours diversification of chromosomal rearrangements that creates and maintains association among locally adapted loci (Kirkpatrick & Barton 2006).

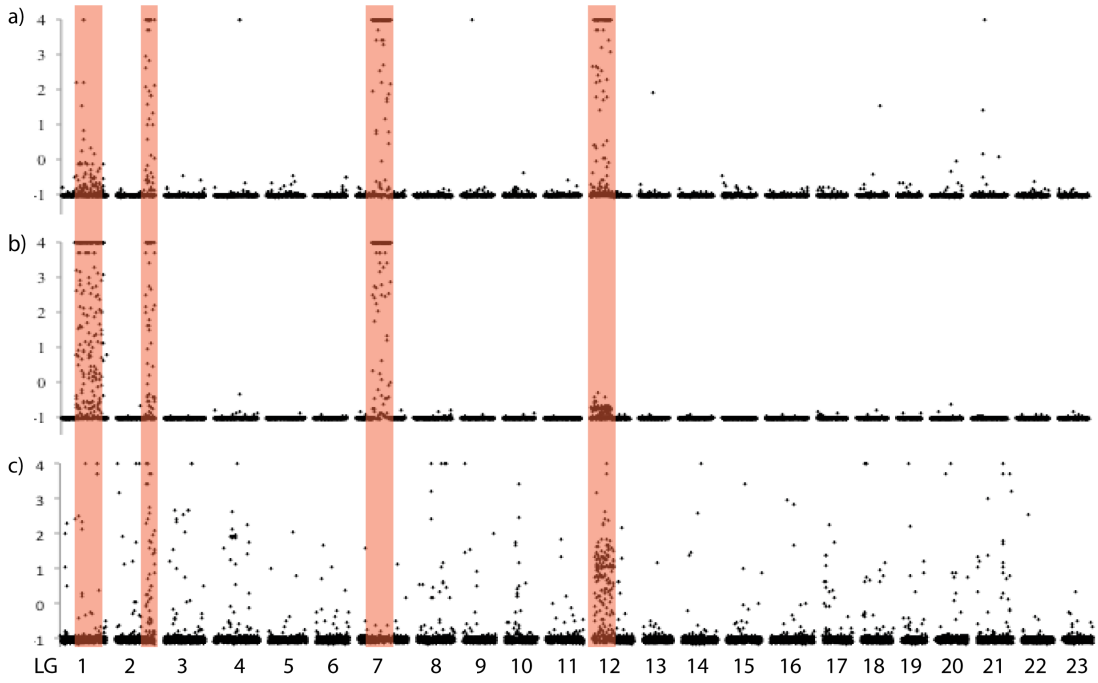

Figure 3. Manhattan plots, visualizing highly divergent genomic regions in Atlantic cod. The outlier pattern between (**a**) northern and southern cod populations in the Northwest Atlantic, (**b**) migratory and non-migratory populations/ecotypes in the Northeast Atlantic and (**c**) Baltic and North Sea/Kattegat populations, indicate that the majority of genomic divergence in the Atlantic cod genome are clustered in four large divergent regions. The plots are based on median $\log_{10}$ Posterior Odds (PO) values from 10 independent runs of BAYESCAN.

In line with these findings, we observe generally low genome wide divergence interspersed with highly divergent regions among the investigated Atlantic cod

populations (Fig. 3), where gene flow could potentially be high due to few physical barriers (paper II, III and IV). However, the genome wide divergence between the well-separated Northeast Atlantic, Northwest Atlantic and the Baltic Sea areas can be substantial, which is reflected in the neutral population structure (Fig. 4).

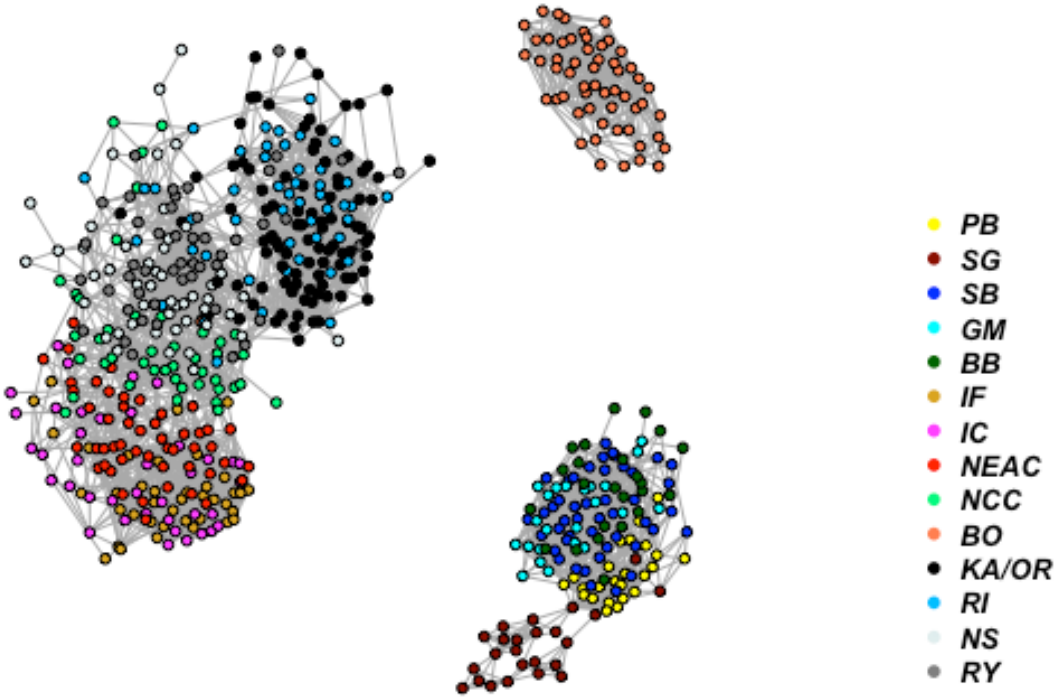

Figure 4. Neutral population divergence between Atlantic cod populations. Population clustering based on 7,075 neutral SNPs in Atlantic cod individuals representing all main areas investigated in this thesis. The analyses are based on an isolation by state (IBS) matrix constructed in PLINK and visualized using the NETVIEW-p pipeline at k=20 that captures the large-scale genetic differentiation between all individuals. Edge width is proportional to the genetic distance between individuals. PB = Placentia Bay, SG = Southern Gulf of St. Lawrence, SB = Sambro, GM = Gulf of Maine, BB = Browns Bank, IF = Iceland Frontal, IC = Iceland Coastal, NEAC = Northeast Arctic cod, NCC = Norwegian coastal cod, BO = Bornholm (representing the Baltic population), KA/OR = Kattegat and Oresund combined, RI = Risør Inner fjord, NS = North Sea, RY = Risør outer fjord.

In an adaptation with gene flow scenario where different morphs or ecotypes are maintained in close proximity and potentially interbreeding, such as in the case of NEAC and NCC (paper II), inversion polymorphism (effectively acting as supergenes) could be an important factor in upholding the morph or ecotype diversity. This has been shown in e.g. *Heliconius* butterflies, where

supergenes controlling wing mimicry has been allocated to a series of inversions that supress recombination (Joron *et al.* 2011; Jones *et al.* 2012b). Chromosomal rearrangements has also recently been associated with behavioural differences in species such as white-throated sparrow (*Zonotrichia albicollis*) (Zinzow-Kramer *et al.* 2015), rainbow trout (*Oncorhynchus mykiss*) (Pearse *et al.* 2014) and fire ant (*Solenopsis Invicta*) (Wang *et al.* 2013). In 2010 Lowry and Willis (2010) showed, for the first time, a direct link between local adaptation and the effects of inversions using QTL mapping in a field experiment in the monkeyflower (*Mimulus guttatus*). Subsequently, it has been shown that perennial and annual ecotypes of monkeyflower, differed significantly within an inversion while high gene flow homogenized the collinear parts of the genome (Twyford & Friedman 2015) and candidate genetic mechanisms for this adaptation have recently been investigated (Gould, Chen & Lowry 2016). Pure ecotype divergence, associated with inversions are also known from e.g. marine and freshwater adaptation in stickleback (*Gasterosteus aculeatus*) (Jones *et al.* 2012a), resident and anadromous ecotypes in rainbow trout (*Oncorhynchus mykiss*) (Pearse *et al.* 2014) and ecological isolation in *Anopheles* mosquitoes (Love *et al.* 2016). As such, the differentiation that we observe between the Icelandic coastal and the Icelandic frontal samples (paper IV), which has been grouped according to behaviour (based on DST-tags), probably reflects true ecotype divergence in a similar way.

The recent development of SNP genotyping and high-throughput sequencing technologies have made individual based large-scale genotypic data easily available. In Atlantic cod, the existence of a well annotated genome (Star *et al.* 2011; Tørresen *et al.* 2016), enables us to face many of the questions and

challenges addressed in the earlier population genetic studies on Atlantic cod. We are now starting to explore, in more detail than previously possible, the causes and consequences of genomic adaptation and the footprints of natural selection across a wide range of biotic and abiotic conditions. Hence, we were able to challenge the functional importance of the observed genomic variation coupled to environmental conditions and behavioral differences. This process was initiated by exploring SNPs at relatively high density (Berg *et al.* 2015 (paper I); 2016 (paper II); Sodeland *et al.* 2016 (paper III); Kirubakaran *et al.* 2016), as part of the Cod SNP Consortium and will continue in future projects, based on next-generation sequencing in e.g. the Aqua Genome Project (www.aquagenome.uio.no). Hence, important questions will be (and are being) asked and it is likely to change our understanding of the biology of Atlantic cod and other teleost species.

FUTURE DIRECTIONS

The genomic architecture, the genetic mechanisms that facilitates adaptive radiation, and how environmental factors can lead to genetic adaptation are keys to an understanding of how biological diversity arises and these issues have focal points in current evolutionary biology (Kawecki & Ebert 2004). SNPs have been the preferred choice of genetic/genomic markers for ecologists and evolutionary biologists for more than a decade, but this is now changing to large scale re-sequencing and *de novo* sequencing of full genomes, due to reduced sequencing costs and technological advances – including bioinformatics analyses. Hence, the numbers of sequenced genomes from non-model organisms are accumulating (Ellegren 2013). It is likely that progress in analyses of non-model organisms will be made on the background of methodology and research tools developed within human research and established model organisms. Along with the added value that "the sequencing revolution" constitute to population genetic research, key taxa can be sequenced to explore long-standing phylogenetic questions and controversies. Phylogenetic analyses based on fully sequenced genomes can provide increased resolution and accuracy to pinpoint the origin of specific adaptive traits (see e.g. Malmstrøm *et al.* 2016), yielding valuable insight to the link between genotype and phenotype at an evolutionary level.

So far, the majority of all re-sequencing initiatives have been based on short read technology (Illumina, SOLiD, 454 and Ion Torrent), which has its limitations in sufficiently covering long repetitive elements, copy number alterations and structural variation (Goodwin, McPherson & McCombie 2016). By using long read sequencing technologies (read lengths of several kb) such as

PacBio and Oxford Nanopore, it is possible to read through large repetitive elements and reveal long-range genomic structures and splice variants in a single continuous read. The improved read length also implies an increased ability to phase chromosomes (Kuleshov *et al.* 2014), which in turn enables tracking of genomic variation across generations. By using long read sequencing technology, state of the art *de novo* genome assemblies can be generated, which is especially useful when sequencing non-model organisms and to improve existing genome assemblies. Novel opportunities also arise from tiny nanopore sequencers as these could be brought into the field where sequencing and data analyses could be performed on site (Quick *et al.* 2016). Even though these long read technologies have lower throughput and are currently more expensive than short read technology, they are foreseen to have an increased role in future sequencing projects and the advantages within non-model species could (and probably will) be substantial.

Another promising outcome of the sequencing revolution is the impact it has on epigenetic research, not only in human systems. Screening of DNA methylation patterns could also have large implications for our understanding of local adaptation as epigenetic modifications are expected to occur at a much faster rate than genetic mutations, which could cause high levels of genome-wide divergence between individuals occupying different habitats (Trucchi *et al.* 2016). As such, population profiling based DNA methylation patterns is likely to give valuable insight into the adaptive role of gene regulation. Simultaneously, such analyses can shed light on plastic responses to different environments, mediated by either traditional gene regulation or epigenetics.

Given that large regions under selection are detected, either by SNP

genotyping or by NGS sequencing, it is important to get an ultimate understanding of the actual targets of selection within these regions. Hence, developing novel methods aiming at identifying the targets of selection within tightly linked genomic regions, where traditional methods have limited powers, are of utmost importance. One possible approach is to do functional testing of candidate genes, but this requires genetic manipulation and the creation of transgenic or knockout animals, tools that are not readily available in non-model organisms such as Atlantic cod. However, the advent of the CRISPR-Cas9 genome editing system offers a promising tool to unravel the causality between genes, phenotypes and fitness in both model and non-model organisms (Bono, Olesnicky & Matzkin 2015), particularly for traits that have a relatively simple genomic architecture (Wolf & Ellegren 2016). In theory, the CRISPR-Cas9 methodology provides an easy and accessible framework for genome editing. The methodology can be used to knock out (or in) specific candidate genes or to make transgenic animals with precise functional (protein) modifications as well as gene regulation changes. Noteworthy, there are practical challenges in incorporating the CRISPR-Cas9 system at early stage embryos in non-model species, but the system has been implemented in zebrafish (e.g. Hwang *et al.* 2013) and recently also successfully in Atlantic salmon (Edvardsen *et al.* 2013; Wargelius *et al.* 2016) and tilapia (Li *et al.* 2014). Given that the initial challenges are overcome, the use of the CRISPR-Cas9 methodology holds great promise in identifying the underlying genetic basis of adaptation and divergence in Atlantic cod and other non-model species, as well as for a deeper understanding of evolutionary biology in general.

ACKNOWLEDGEMENTS

I would like to thank Helle T. Baalsrud, Martin Malmstøm, Julia M. I. Barth and Kjetil L. Voje for useful comments and proofreading of the thesis.

REFERENCES

Andersen, O. (2012) Hemoglobin polymorphisms in Atlantic cod - A review of 50years of study. *Marine genomics*, **8**, 59–65.

Andersen, O., Johnsen, H., De Rosa, M.C., Præbel, K., Stjelja, S., Kirubakaran, T.G., Pirolli, D., Jentoft, S. & Fevolden, S.-E. (2015) Evolutionary history and adaptive significance of the polymorphic Pan I in migratory and stationary populations of Atlantic cod (*Gadus morhua*). *Marine genomics*, **22**, 45–54.

Andersen, O., Wetten, O.F., De Rosa, M.C., André, C., Alinovi, C.C., Colafranceschi, M., Brix, O. & Colosimo, A. (2009) Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 833–841.

André, C., Larsson, L.C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G.R., Dahlgren, T.G., Hutchinson, W.F., Mariani, S., Mudde, K., Ruzzante, D.E. & Ryman, N. (2011) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. *Heredity*, **106**, 270–280.

Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A. & Luikart, G. (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC bioinformatics*, **9**:323.

Ayala, F.J. & Coluzzi, M. (2005) Chromosome speciation: humans, *Drosophila*, and mosquitoes. *Proceedings of the National Academy of Sciences of the United States of America*, **102 Suppl 1**, 6535–6542.

Árnason, E. (2004) Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: Trans-Atlantic clines and shallow gene genealogy. *Genetics*, **166**, 1871–1885.

Árnason, E. & Palsson, S. (1996) Mitochondrial cytochrome *b* DNA sequence variation of Atlantic cod *Gadus morhua*, from Norway. *Molecular Ecology*, **5**, 715–724.

Árnason, E., Hernandez, U.B. & Kristinsson, K. (2009) Intense habitat-specific fisheries-induced selection at the molecular *Pan* I locus predicts imminent collapse of a major cod fishery. *Plos One*, **4**:e5529.

Árnason, E., Palsson, S. & Arason, A. (1992) Gene Flow and Lack of Population Differentiation in Atlantic Cod, *Gadus-Morhua* L., From Iceland, and Comparison of Cod From Norway and Newfoundland. *Journal of Fish Biology*, **40**, 751–770.

Árnason, E., Petersen, P.H. & Palsson, S. (1998) Mitochondrial cytochrome *b* DNA sequence variation of Atlantic cod, *Gadus morhua*, from the Baltic and the White Seas. *Hereditas*, **129**, 37–43.

Barth, J.M.I., Berg, P.R., Jonsson, P.R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., Jacobsen, K.S., Johannesson, K., Jorde, P.E., Knutsen, H., Moksnes, P-O., Star, B., Stenseth, N.C., Svedäng, H., Jentoft, S., & André, C. Local adaptation of Atlantic cod (*Gadus morhua*) to fjord ecosystems despite high connectivity. Manuscript in prep.

Beacham, T.D., Brattey, J., Miller, K.M., Le, K.D. & Withler, R.E. (2002) Multiple stock structure of Atlantic cod (*Gadus morhua*) off Newfoundland and Labrador determined from genetic variation. *ICES Journal of Marine Science: Journal du Conseil*, **59**, 650–665.

Beaumont, M.A. & Nichols, R.A. (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.

Begg, G.A. & Marteinsdottir, G. (2000) Spawning origins of pelagic juvenile cod *Gadus morhua* inferred from spatially explicit age distributions: potential influences on year-class strength and recruitment. *Marine Ecology Progress Series*, **202**, 193–217.

Bentzen, P., Taggart, C.T., Ruzzante, D.E. & Cook, D. (1996) Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the northwest Atlantic. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 2706–2721.

Berg, P.R., Jentoft, S., Star, B., Ring, K.H., Knutsen, H., Lien, S., Jakobsen, K.S. & André, C. (2015) Adaptation to Low Salinity Promotes Genomic Divergence in Atlantic Cod (*Gadus morhua* L.). *Genome biology and evolution*, **7**, 1644–1663.

Berg, P.R., Star, B., Pampoulie, C., Sodeland, M., Barth, J.M.I., Knutsen, H., Jakobsen, K.S. & Jentoft, S. (2016) Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific reports*, **6**:23246.

Bergstad, O.A., Jørgensen, T. & Dragesund, O. (1987) Life history and ecology of the gadoid resources of the Barents Sea. *Fisheries Research*, **5**, 119–161.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011). The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* **20**: 2044-2072.

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S. & Sancristobal, M. (2010) Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, **186**, 241–262.

Bono, J.M., Olesnicky, E.C. & Matzkin, L.M. (2015) Connecting genotypes, phenotypes and fitness: harnessing the power of CRISPR/Cas9 genome editing. *Molecular Ecology*, **24**, 3810–3822.

Borza, T., Stone, C., Gamperl, A.K. & Bowman, S. (2009) Atlantic cod (*Gadus*

morhua) hemoglobin genes: multiplicity and polymorphism. *BMC genetics*, **10**:51.

Bradbury, I.R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I.G., Snelgrove, P.V.R., Morris, C.J., Gregory, R.S., Hardie, D.C., Hutchings, J.A., Ruzzante, D.E., Taggart, C.T. & Bentzen, P. (2010) Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3725–3734.

Bradbury, I.R., Hubert, S., Higgins, B., Bowman, S., Borza, T., Paterson, I.G., Snelgrove, P.V.R., Morris, C.J., Gregory, R.S., Hardie, D., Hutchings, J.A., Ruzzante, D.E., Taggart, C.T. & Bentzen, P. (2013) Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evolutionary Applications*, **6**, 450–461.

Brander, K. (2005) Spawning and life history information for North Atlantic cod stocks. *ICES Cooperative Research Report*, **274**, 1–152.

Brix, O., Thorkildsen, S. & Colosimo, A. (2004) Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes *HbI*1/1, HbI*1/2, and HbI*2/2* by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). *Comparative Biochemistry and Physiology Part A*, **138**, 241–251.

Caddy, J.F. & Cochrane, K.L. (2001) A review of fisheries management past and present and some future perspectives for the third millennium. *Ocean and Coastal Management*, **44**, 653–682.

Cano, J.M., Shikano, T., Kuparinen, A. & Merila, J. (2008) Genetic differentiation, effective population size and gene flow in marine fishes: implications for stock management. *Journal of Integrative Field Biology*, **5**, 1–10.

Carr, S.M. & Marshall, H.D. (1991) Detection of Intraspecific.DNA Sequence Variation in the Mitochondrial Cytochrome *b* Gene of Atlantic Cod (*Gadus morhua*) by the Polymerase Chain Reaction. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 48–52.

Carr, S.M., Snellen, A.J., Howse, K.A. & Wroblewski, J.S. (1995) Mitochondrial DNA sequence variation and genetic stock structure of Atlantic cod (*Gadus morhua*) from bay and offshore locations on the Newfoundland continental shelf. *Molecular Ecology*, **4**, 79–88.

Case, R., Hutchinson, W., Hauser, L., Van Oosterhout, C. & Carvalho, G. (2005) Macro- and micro-geographic variation in pantophysin (*Pan*I) allele frequencies in NE Atlantic cod *Gadus morhua*. *Marine Ecology Progress Series*, **301**, 267–278.

Charlesworth, B., Nordborg, M. & Charlesworth, D. (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium

patterns of genetic diversity in subdivided populations. *Genetical research*, **70**, 155–174.

Cheng, C., White, B.J., Kamdem, C., Mockaitis, K., Costantini, C., Hahn, M.W. & Besansky, N.J. (2012) Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics*, **190**, 1417–1432.

Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J.K. (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.

Cutter, A.D. (2006) Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer Caenorhabditis elegans. *Genetics*, **172**, 171–184.

de Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J. & Vigouroux, Y. (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

de Villemereuil, P., Frichot, É., Bazin, E., François, O. & Gaggiotti, O.E. (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

De Vos, J.M., Joppa, L.N., Gittleman, J.L., Stephens, P.R. & Pimm, S.L. (2015) Estimating the normal background rate of species extinction. *Conservation Biology*, **29**, 452–462.

Deagle, B.E., Jones, F.C., Absher, D.M., Kingsley, D.M. & Reimchen, T.E. (2013) Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. *Molecular Ecology*, **22**, 1917–1932.

Defaveri, J., Shikano, T., Shimada, Y. & Merilä, J. (2013) High degree of genetic differentiation in marine three-spined sticklebacks (*Gasterosteus aculeatus*). *Molecular Ecology*, **22**, 4811–4828.

Defaveri, J., Shikano, T., Shimada, Y., Goto, A. & Merilä, J. (2011) Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution*, **65**, 1800–1807.

Dobzhansky, T. (1943) Genetics of Natural Populations IX. Temporal Changes in the Composition of Populations of *Drosophila pseudoobscura*. *Genetics*, **28**, 162–186.

Edvardsen, R.B., Leininger, S., Kleppe, L., Skaftnesmo, K.O. & Wargelius, A. (2013) Targeted mutagenesis in Atlantic salmon (*Salmo salar* L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation. *Plos One*, **9**:e108622.

Ellegren, H. (2013) Genome sequencing and population genomics in non-model organisms. *Trends in ecology & evolution*, **29**, 51–63.

Ellegren, H. & Sheldon, B.C. (2008) Genetic basis of fitness differences in natural populations. *Nature*, **452**, 169–175.

Feder, J.L., Egan, S.P. & Nosil, P. (2012a) The genomics of speciation-with-gene-flow. *Trends in Genetics*, **28**, 342–350.

Feder, J.L., Gejji, R., Yeaman, S. & Nosil, P. (2012b) Establishment of new mutations under divergence and genome hitchhiking. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, **367**, 461–474.

Feder, J.L., Nosil, P. & Flaxman, S.M. (2014) Assessing when chromosomal rearrangements affect the dynamics of speciation: implications from computer simulations. *Frontiers in Genetics*, **5**:295.

Feulner, P.G.D., Chain, F.J.J., Panchal, M., Eizaguirre, C., Kalbe, M., Lenz, T.L., Mundry, M., Samonte, I.E., Stoll, M., Milinski, M., Reusch, T.B.H. & Bornberg-Bauer, E. (2013) Genome-wide patterns of standing genetic variation in a marine population of three-spined sticklebacks. *Molecular Ecology*, **22**, 635–649.

Fevolden, S.-E. & Pogson, G.H. (1996) A highly diagnostic nuclear DNA Marker for distinguishing between Norwegian coastal and Northeast Arctic populations of Atlantic cod. *ICES C.M.* 1996/G: **24**.

Fevolden, S.-E. & Pogson, G.H. (1997) Genetic divergence at the synaptophysin (*Syp* I) locus among Norwegian coastal and north-east Arctic populations of Atlantic cod. *Journal of Fish Biology*, **51**, 895–908.

Fevolden, S.-E., Westgaard, J.I., Pedersen, T. & Præbel, K. (2012) Settling-depth vs. genotype and size vs. genotype correlations at the *Pan* I locus in 0-group Atlantic cod *Gadus morhua*. *Marine Ecology Progress Series*, **468**, 267–278.

Foll, M. & Gaggiotti, O. (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Frydenberg, O., Møller, D., Naevdal, G. & Sick, K. (1965) Haemoglobin polymorphism in Norwegian cod populations. *Hereditas*, **53**, 257–271.

Fuller, Z.L., Haynes, G.D., Richards, S. & Schaeffer, S.W. (2016) Genomics of Natural Populations: How Differentially Expressed Genes Shape the Evolution of Chromosomal Inversions in *Drosophila pseudoobscura*. *Genetics*, **204**, 287–301.

Funk, W.C., McKay, J.K., Hohenlohe, P.A. & Allendorf, F.W. (2012) Harnessing genomics for delineating conservation units. *Trends in ecology & evolution*, **27**, 489–496.

Gaggiotti, O.E., Bekkevold, D., Jørgensen, H.B.H., Foll, M., Carvalho, G.R., André, C. & Ruzzante, D.E. (2009) Disentangling the Effects of Evolutionary, Demographic, and Environmental Factors Influencing Genetic Structure of Natural Populations: Atlantic Herring as a Case Study. *Evolution*, **63**, 2939–2951.

Gagnaire, P.-A., Pavey, S.A., Normandeau, E. & Bernatchez, L. (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483–2497.

Gewin, V. (2015) North Pacific "blob" stirs up fisheries management. *Nature*, **524**, 396.

Goodwin, S., McPherson, J.D. & McCombie, W.R. (2016) Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, **17**, 333–351.

Gould, B.A., Chen, Y. & Lowry, D.B. (2016) Pooled Ecotype Sequencing Reveals Candidate Genetic Mechanisms for Adaptive Differentiation and Reproductive Isolation. *Molecular Ecology*, early view, doi: 10.1111/mec.13881.

Guerrero, R.F., Rousset, F. & Kirkpatrick, M. (2012) Coalescent patterns for chromosomal inversions in divergent populations. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **367**, 430–438.

Guinand, B., Lemaire, C. & Bonhomme, F. (2004) How to detect polymorphisms undergoing selection in marine fishes? A review of methods and case studies, including flatfishes. *Journal of Sea Research*, **51**, 167–182.

Guo, B., Defaveri, J., Sotelo, G., Nair, A. & Merilä, J. (2014) Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, **13**:19.

Guo, B., Li, Z. & Merilä, J. (2016) Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Molecular Ecology*, **25**, 2833–2852.

Günther, T. & Coop, G. (2013) Robust identification of local adaptation from allele frequencies. *Genetics*, **195**, 205–220.

Gyllensten, U. (1985) The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous, and freshwater species. *Journal of Fish Biology*, **26**, 691–699.

Hardie, D.C. & Hutchings, J.A. (2011) The Ecology of Atlantic Cod (*Gadus morhua*) in Canadian Arctic Lakes. *Arctic*, **64**, 137–150.

Hardie, D.C., Gillett, R.M. & Hutchings, J.A. (2006) The effects of isolation and colonization history on the genetic structure of marine-relict populations of Atlantic cod (*Gadus morhua*) in the Canadian Arctic. *Canadian Journal of*

Fisheries and Aquatic Sciences, **63**, 1830–1839.

Harrald, M., Wright, P.J. & Neat, F.C. (2010) Substock variation in reproductive traits in North Sea cod (*Gadus morhua*). *Canadian Journal of Fisheries and Aquatic Sciences*, **67**, 866–876.

Hartl, D.L. & Clark, A.G. (2007) *Principles of Population Genetics*. Sinauer Associates Incorporated.

Hauser, L. & Carvalho, G.R. (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, **9**, 333–362.

Hemmer-Hansen, J., Nielsen, E.E., Frydenberg, J. & Loeschcke, V. (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). *Heredity*, **99**, 592–600.

Hemmer-Hansen, J., Nielsen, E.E., Therkildsen, N.O., Taylor, M.I., Ogden, R., Geffen, A.J., Bekkevold, D., Helyar, S., Pampoulie, C., Johansen, T., FishPopTrace consortium & Carvalho, G.R. (2013) A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*, **22**, 2653–2667.

Hendry, A.P., Nosil, P. & Rieseberg, L.H. (2007) The speed of ecological speciation. *Functional ecology*, **21**, 455–464.

Hill, W.G. & Robertson, A. (1968) Linkage disequilibrium in finite populations. *Theoretical and applied genetics*, **38**, 226–231.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. & Cresko, W.A. (2010) Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLoS Genetics*, **6**:e1000862.

Hooper, D.M. & Price, T.D. (2015) Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution*, **69**, 890–903.

Howell, W.H., Morin, M., Rennels, N. & Goethel, D. (2008) Residency of adult Atlantic cod (*Gadus morhua*) in the western Gulf of Maine. *Fisheries Research*, **91**, 123–132.

Hubert, S., Higgins, B., Borza, T. & Bowman, S. (2010) Development of a SNP resource and a genetic linkage map for Atlantic cod (*Gadus morhua*). *BMC genomics*, **11**:191.

Hutchings, J.A. (2000) Collapse and recovery of marine fishes. *Nature*, **406**, 882–885.

Hutchings, J.A., Swain, D.P., Rowe, S., Eddington, J.D., Puvanendran, V. & Brown, J.A. (2007) Genetic Variation in Life-History Reaction Norms in a Marine Fish. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1693–1699.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J. & Joung, J.K. (2013) Efficient genome editing in zebrafish using a

CRISPR-Cas system. *Nature Biotechnology*, **31**, 227–229.

Hylen, A. (1964) Coastal Cod and Skrei in the Lofoten Area. *Fiskeridirektoratets Skrifter Serie Havundersøkelser*, **13**, 27–42.

Imsland, A., Foss, A., Naevdal, G., Johansen, T., Folkvord, A., Stefansson, S. & Jonassen, T. (2004) Variations in growth in haemoglobin genotypes of Atlantic cod. *Fish Physiology and Biochemistry*, **30**, 47–55.

Jacobsen, M.W., Pujolar, J.M., Bernatchez, L., Munch, K., Jian, J., Niu, Y. & Hansen, M.M. (2014) Genomic footprints of speciation in Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Molecular Ecology*, **23**, 4785–4798.

Jakobsdóttir, K.B., Pardoe, H., Magnússon, Á., Björnsson, H., Pampoulie, C., Ruzzante, D.E. & Marteinsdóttir, G. (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**, 562–573.

Jakobsen, T. (1987) Coastal cod in northern Norway. *Fisheries Research*, **5**, 223–234.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M.C., Myers, R.M., Miller, C.T., Summers, B.R., Knecht, A.K., Brady, S.D., Zhang, H., Pollen, A.A., Howes, T., Amemiya, C., Baldwin, J., Bloom, T., Jaffe, D.B., Nicol, R., Wilkinson, J., Lander, E.S., Di Palma, F., Lindblad-Toh, K. & Kingsley, D.M. (2012a) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Jones, R.T., Salazar, P.A., ffrench-Constant, R.H., Jiggins, C.D. & Joron, M. (2012b) Evolution of a mimicry supergene from a multilocus architecture. *Proceedings: Biological Sciences*, **279**, 316–325.

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., Wilkinson, P.A., Salazar, C., Davidson, C., Clark, R., Quail, M.A., Beasley, H., Glithero, R., Lloyd, C., Sims, S., Jones, M.C., Rogers, J., Jiggins, C.D. & ffrench-Constant, R.H. (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*, **477**, 203–206.

Jónsson, S. & Valdimarsson, H. (2005) The flow of Atlantic water to the North Icelandic Shelf and its relation to the drift of cod larvae. *ICES Journal of Marine Science: Journal du Conseil*, **62**, 1350–1359.

Jørgensen, T. (1992) Long-term changes in growth of North-east Arctic cod (*Gadus morhua*) and some environmental influences. *ICES Journal of Marine Science: Journal du Conseil*, **49**, 263–278.

Kandul, N.P., Lukhtanov, V.A. & Pierce, N.E. (2007) Karyotypic diversity and speciation in *Agrodiaetus* butterflies. *Evolution*, **61**, 546–559.

Karlsen, B.O., Klingan, K., Emblem, A., Jørgensen, T.E., Jueterbock, A., Furmanek, T., Hoarau, G., Johansen, S.D., Nordeide, J.T. & Moum, T. (2013) Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology*, **22**, 5098–5111.

Karlsson, S. & Mork, J. (2003) Selection-induced variation at the pantophysin locus (PanI) in a Norwegian fjord population of cod (*Gadus morhua L.*). *Molecular Ecology*, **12**, 3265–3274.

Karpov, A. K. & Novikov, G. G. (1980) Hemoglobin alloforms in cod, *Gadus morhua* (Gadiformes, Gadidae), their functional characteristics and occurrence in populations. *Journal of Ichthyology*, **6**, 45-49.

Kawecki, T.J. & Ebert, D. (2004) Conceptual issues in local adaptation. *Ecology letters*, **7**, 1225–1241.

Kimura, M. (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*, **61**, 893-903.

Kimura, M. & Crow, J.F. (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725–738.

Kirkpatrick, M. & Barton, N. (2006) Chromosome inversions, local adaptation and speciation. *Genetics*, **173**, 419–434.

Kirkpatrick, M. & Kern, A. (2012) Where's the money? Inversions, genes, and the hunt for genomic targets of selection. *Genetics*, **190**, 1153–1155.

Kirubakaran, T.G., Grove, H., Kent, M.P., Sandve, S.R., Baranski, M., Nome, T., De Rosa, M.C., Righino, B., Johansen, T., Otterå, H., Sonesson, A., Lien, S. & Andersen, O. (2016) Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology*, **25**, 2130–2143.

Knutsen, H., Jorde, P.E., André, C. & Stenseth, N.C. (2003) Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology*, **12**, 385–394.

Kuleshov, V., Xie, D., Chen, R., Pushkarev, D., Ma, Z., Blauwkamp, T., Kertesz, M. & Snyder, M. (2014) Whole-genome haplotyping using long reads and statistical methods. *Nature Biotechnology*, **32**, 261–266.

Kurlansky, M. (1997) *Cod: a Biography of the Fish That Changed the World*. Walker and Company, New York.

Laikre, L., Palm, S. & Ryman, N. (2005) Genetic Population Structure of Fishes: Implications for Coastal Zone Management. *AMBIO: A Journal of the Human Environment*, **34**, 111–119.

Lamichhaney, S., Martinez Barrio, A., Rafati, N., Sundström, G., Rubin, C.-J., Gilbert, E.R., Berglund, J., Wetterbom, A., Laikre, L., Webster, M.T., Grabherr, M.,

Ryman, N. & Andersson, L. (2012) Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 19345–19350.

Laporte, M., Rogers, S.M., Dion-Cote, A.M., Normandeau, E. & Gagnaire, P.A. (2015) RAD-QTL Mapping Reveals Both Genome-Level Parallelism and Different Genetic Architecture Underlying the Evolution of Body Shape in Lake Whitefish (*Coregonus clupeaformis*) Species Pairs. *G3*, **5**, 2919–2919.

Lewontin, R.C. & Kojima, K.-I. (1960) The Evolutionary Dynamics of Complex Polymorphisms. *Evolution*, **14**, 458–472.

Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Li, M., Yang, H., Zhao, J., Fang, L., Shi, H., Li, M., Sun, Y., Zhang, X., Jiang, D., Zhou, L. & Wang, D. (2014) Efficient and Heritable Gene Targeting in Tilapia by CRISPR/Cas9. *Genetics*, **197**, 591–599.

Lotterhos, K.E. & Whitlock, M.C. (2014) Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Love, R.R., Steele, A.M., Coulibaly, M.B., Traore, S.F., Emrich, S.J., Fontaine, M.C. & Besansky, N.J. (2016) Chromosomal inversions and ecotypic differentiation in *Anopheles gambiae*: the perspective from whole-genome sequencing. *Molecular Ecology*, **25**, 5889–5906.

Lowry, D.B. & Willis, J.H. (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology*, **8**:e1000500.

Luikart, G., England, P.R., Tallmon, D., Jordan, S. & Taberlet, P. (2003) The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Mack, K.L., Campbell, P. & Nachman, M.W. (2016) Gene regulation and speciation in house mice. *Genome Research*, **26**, 451–461.

Makeenko, G.A., Volkov, A.A., Mugue, N.S. & Zelenina, D.A. (2014) Genetic variation at the pantophysin (*Pan*I) locus in North-East Arctic cod *Gadus morhua* L. (Godiformes: Gadidae) population in the Barents Sea and adjacent waters. *Russian Journal of Genetics*, **50**, 1271–1284.

Malmstrøm, M., Matschiner, M., Tørresen, O.K., Star, B., Snipen, L.G., Hansen, T.F.,

Baalsrud, H.T., Nederbragt, A.J., Hanel, R., Salzburger, W., Stenseth, N.C., Jakobsen, K.S. & Jentoft, S. (2016) Evolution of the immune system influences speciation rates in teleost fishes. *Nature Genetics*, **48**, 1204–1210.

Marteinsdottir, G. & Begg, G.A. (2002) Essential relationships incorporating the influence of age, size and condition on variables required for estimation of reproductive potential in Atlantic cod *Gadus morhua*. *Marine Ecology Progress Series*, **235**, 235–256.

Martinez, P.A., Zurano, J.P., Amado, T.F., Penone, C., Betancur-R, R., Bidau, C.J. & Jacobina, U.P. (2015) Chromosomal diversity in tropical reef fishes is related to body size and depth range. *Molecular phylogenetics and evolution*, **93**, 1–4.

Matzkin, L.M., Merritt, T.J.S., Zhu, C.-T. & Eanes, W.F. (2005) The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion *In(3R)Payne* in *Drosophila melanogaster*. *Genetics*, **170**, 1143–1152.

Mendel, G. (1866) *Versuche Über Pflanzenhybriden*. Verhandlungen des naturforschenden Vereines in Brünn.

Michalsen, K., Johansen, T., Subbey, S. & Beck, A. (2014) Linking tagging technology and molecular genetics to gain insight in the spatial dynamics of two stocks of cod in Northeast Atlantic waters. **71**, 1417–1432.

Mieszkowska, N., Genner, M.J., Hawkins, S.J. & Sims, D.W. (2009) *Effects of Climate Change and Commercial Fishing on Atlantic Cod*, 1st ed. Elsevier Ltd.

Moen, T., Ben Hayes, Nilsen, F., Delghandi, M., Fjalestad, K.T., Fevolden, S.-E., Berg, P.R. & Lien, S. (2007) Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. *BMC genetics*, **9**:18.

Moller, D. (1968) Genetic diversity in spawning cod along the Norwegian coast. *Hereditas*, **60**, 1–32.

Moritz, C. (1994) Defining Evolutionarily-Significant-Units for Conservation. *Trends in ecology & evolution*, **9**, 373–375.

Mork, J., Ryman, N., Ståhl, G., Utter, F. & Sundnes, G. (1985) Genetic Variation in Atlantic Cod (*Gadus morhua*) Throughout Its Range. *Canadian Journal of Fisheries and Aquatic Sciences*, **42**, 1580–1587.

Nadeau, N.J., Whibley, A., Jones, R.T., Davey, J.W., Dasmahapatra, K.K., Baxter, S.W., Quail, M.A., Joron, M., ffrench-Constant, R.H., Blaxter, M.L., Mallet, J. & Jiggins, C.D. (2012) Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **367**, 343–353.

Narum, S.R. & Hess, J.E. (2011) Comparison of F_{ST} outlier tests for SNP loci under

selection. *Molecular Ecology Resources*, **11 Suppl 1**, 184–194.

Nielsen, E.E., Cariani, A., Mac Aoidh, E., Maes, G.E., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M.T., Martinsohn, J.T., McEwing, R., Panitz, F., Patarnello, T., Tinti, F., Van Houdt, J.K.J., Volckaert, F.A.M., Waples, R.S., FishPopTrace consortium, Albin, J.E.J., Vieites Baptista, J.M., Barmintsev, V., Bautista, J.M., Bendixen, C., Bergé, J.-P., Blohm, D., Cardazzo, B., Diez, A., Espiñeira, M., Geffen, A.J., Gonzalez, E., González-Lavín, N., Guarniero, I., Jeráme, M., Kochzius, M., Krey, G., Mouchel, O., Negrisolo, E., Piccinetti, C., Puyet, A., Rastorguev, S., Smith, J.P., Trentini, M., Verrez-Bagnis, V., Volkov, A., Zanzi, A. & Carvalho, G.R. (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature communications*, **3**:851.

Nielsen, E.E., Hansen, M.M. & Meldrup, D. (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology*, **15**, 3219–3229.

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

Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F. & Bekkevold, D. (2009a) Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology*, **18**, 3128–3150.

Nielsen, E.E., Hemmer-Hansen, J., Poulsen, N.A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G.-L., Ogden, R. & Carvalho, G.R. (2009b) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, **9**:276.

Nielsen, E.E., Wright, P.J., Hemmer-Hansen, J., Poulsen, N.A., Gibb, I.M. & Meldrup, D. (2009c) Microgeographical population structure of cod *Gadus morhua* in the North Sea and west of Scotland: The role of sampling loci and individuals. *Marine Ecology Progress Series*, **376**, 213–225.

Noor, M.A., Grams, K.L., Bertucci, L.A. & Reiland, J. (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 12084–12088.

Nordeide, J. (1998) Coastal cod and north-east Arctic cod - Do they mingle at the spawning grounds in Lofoten? *Sarsia*, **83**, 373–379.

Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

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., Carlson, S.M., Gjøsaeter, J. & Stenseth, N.C. (2009) Nine decades of decreasing phenotypic variability in Atlantic cod. *Ecology letters*, **12**, 622–631.

Olsen, E.M., Heino, M., Lilly, G.R., Morgan, M.J., Brattey, J., Ernande, B. & Dieckmann, U. (2004) Maturation trends indicative of rapid evolution preceded the collapse of northern cod. *Nature*, **428**, 932–935.

Olsen, E.M., Lilly, G.R., Heino, M., Morgan, M.J., Brattey, J. & Dieckmann, U. (2005) Assessing changes in age and size at maturation in collapsing populations of Atlantic cod (Gadus morhua). *Canadian Journal of Fisheries and Aquatic Sciences*, **62**, 811–823.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. & De Meester, L. (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–5999.

Ottersen, G., Michalsen, K. & Nakken, O. (1998) Ambient temperature and distribution of north-east Arctic cod. *ICES journal of marine science*, **55**, 67–85.

Pampoulie, C., Ruzzante, D.E., Chosson, V., Jörundsdóttir, T.D., Taylor, L., Thorsteinsson, V., Daníelsdóttir, A.K. & Marteinsdóttir, G. (2006) The genetic structure of Atlantic cod (G*adus morhua)* around Iceland: insight from microsatellites, the *Pan* I locus, and tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences*, **63**, 2660–2674.

Pampoulie, C., Skirnisdottir, S., Star, B., Jentoft, S., Jonsdottir, I.G., Hjörleifsson, E., Thorsteinsson, V., Pálsson, Ó.K., Berg, P.R., Andersen, O., Magnusdottir, S., Helyar, S.J. & Danielsdottir, A.K. (2015) Rhodopsin gene polymorphism associated with divergent light environments in Atlantic cod. *Behavior Genetics*, **45**, 236–244.

Pearse, D.E., Miller, M.R., Abadía-Cardoso, A. & Garza, J.C. (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, **281**:20140012.

Pershing, A.J., Alexander, M.A., Hernandez, C.M., Kerr, L.A., Bris, A.L., Mills, K.E., Nye, J.A., Record, N.R., Scannell, H.A., Scott, J.D., Sherwood, G.D. & Thomas, A.C. (2015) Slow adaptation in the face of rapid warming leads to collapse of the Gulf of Maine cod fishery. *Science*, **350**, 809–812.

Pogson, G.H. (2001) Nucleotide polymorphism and natural selection at the pantophysin (*Pan* I) locus in the Atlantic cod, *Gadus morhua* (L.). *Genetics*, **157**, 317–330.

Pogson, G.H. & Mesa, K.A. (2004) Positive Darwinian selection at the pantophysin (*Pan* I) locus in marine gadid fishes. *Molecular Biology and Evolution*, **21**, 65–75.

Puig, M., Cáceres, M. & Ruiz, A. (2004) Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 9013–9018.

Pujolar, J.M., Jacobsen, M.W., Als, T.D., Frydenberg, J., Munch, K., Jónsson, B., Jian, J.B., Cheng, L., Maes, G.E., Bernatchez, L. & Hansen, M.M. (2014) Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, **23**, 2514–2528.

Quick, J., Loman, N.J., Duraffour, S., Simpson, J.T., Severi, E., Cowley, L., Bore, J.A., Koundouno, R., Dudas, G., Mikhail, A., Ouédraogo, N., Afrough, B., Bah, A., Baum, J.H.J., Becker-Ziaja, B., Boettcher, J.P., Cabeza-Cabrerizo, M., Camino-Sánchez, Á., Carter, L.L., Doerrbecker, J., Enkirch, T., García-Dorival, I., Hetzelt, N., Hinzmann, J., Holm, T., Kafetzopoulou, L.E., Koropogui, M., Kosgey, A., Kuisma, E., Logue, C.H., Mazzarelli, A., Meisel, S., Mertens, M., Michel, J., Ngabo, D., Nitzsche, K., Pallasch, E., Patrono, L.V., Portmann, J., Repits, J.G., Rickett, N.Y., Sachse, A., Singethan, K., Vitoriano, I., Yemanaberhan, R.L., Zekeng, E.G., Racine, T., Bello, A., Sall, A.A., Faye, O., Faye, O., Magassouba, N., Williams, C.V., Amburgey, V., Winona, L., Davis, E., Gerlach, J., Washington, F., Monteil, V., Jourdain, M., Bererd, M., Camara, A., Somlare, H., Camara, A., Gerard, M., Bado, G., Baillet, B., Delaune, D., Nebie, K.Y., Diarra, A., Savane, Y., Pallawo, R.B., Gutierrez, G.J., Milhano, N., Roger, I., Williams, C.J., Yattara, F., Lewandowski, K., Taylor, J., Rachwal, P., Turner, D.J., Pollakis, G., Hiscox, J.A., Matthews, D.A., O'Shea, M.K., Johnston, A.M., Wilson, D., Hutley, E., Smit, E., Di Caro, A., Wölfel, R., Stoecker, K., Fleischmann, E., Gabriel, M., Weller, S.A., Koivogui, L., Diallo, B., Keïta, S., Rambaut, A., Formenty, P., Günther, S. & Carroll, M.W. (2016) Real-time, portable genome sequencing for Ebola surveillance. *Nature*, **530**, 228–232.

Rak, D. & Wieczorek, P. (2012) Variability of temperature and salinity over the last decade in selected regions of the southern Baltic Sea. *Oceanologia*, **54**, 339–354.

Reiss, H., Hoarau, G., Dickey-Collas, M. & Wolff, W.J. (2009) Genetic population structure of marine fish: mismatch between biological and fisheries management units. *Fish and Fisheries*, **10**, 361–395.

Rieseberg, L.H. (2001) Chromosomal rearrangements and speciation. *Trends in ecology & evolution*, **16**, 351–358.

Righton, D.A., Andersen, K.H., Neat, F., Thorsteinsson, V., Steingrund, P., Svedäng, H., Michalsen, K., Hinrichsen, H.-H., Bendall, V., Neuenfeldt, S., Wright, P., Jonsson, P., Huse, G., Van Der Kooij, J., Mosegaard, H., Hüssy, K. & Metcalfe, J. (2010) Thermal niche of Atlantic cod Gadus morhua: Limits, tolerance and optima. *Marine Ecology Progress Series*, **420**, 1–13.

Roesti, M., Gavrilets, S., Hendry, A.P., Salzburger, W. & Berner, D. (2014) The genomic signature of parallel adaptation from shared genetic variation. *Molecular Ecology*, **23**, 3944–3956.

Roesti, M., Hendry, A.P., Salzburger, W. & Berner, D. (2012) Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular Ecology*, **21**, 2852–2862.

Rollefsen, G. (1933) The otoliths of the cod. *Fiskeridirektoratets Skrifter Serie Havundersøkelser*, **IV**, 1–14.

Ruzzante, D.E., Taggart, C.T. & Cook, D. (1998) A nuclear DNA basis for shelf- and bank-scale population structure in northwest Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. *Molecular Ecology*, **7**, 1663–1680.

Ruzzante, D.E., Taggart, C.T., Doyle, R.W. & Cook, D. (2001) Stability in the historical pattern of genetic structure of Newfoundland cod (*Gadus morhua*) despite the catastrophic decline in population size from 1964 to 1994. *Conservation Genetics*, **2**, 257–269.

Salvanes, A.G.V. & Hart, P.J.B. (2011) Is individual variation in competitive performance of reared juvenile cod influenced by haemoglobin genotype? *Sarsia*, **85**, 265–274.

Salvanes, A., Skjæraasen, J.E. & Nilsen, T. (2004) Sub-populations of coastal cod with different behaviour and life-history strategies. *Marine Ecology Progress Series*, **267**, 241–251.

Sarvas, T.H. & Fevolden, S.-E. (2005) Pantophysin (*Pan* I) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic. *Journal of Fish Biology*, **67**, 444–469.

Savolainen, O., Lascoux, M. & Merilä, J. (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–820.

Schaeffer, S.W. (2008) Selection in heterogeneous environments maintains the gene arrangement polymorphism of *Drosophila pseudoobscura*. *Evolution*, **62**, 3082–3099.

Schwander, T., Libbrecht, R. & Keller, L. (2014) Supergenes and Complex Phenotypes. *Current biology,* **24**, R288–R294.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Sætre, G.-P., Bank, C., Brännström, A., Brelsford, A., Clarkson, C.S., Eroukhmanoff, F., Feder, J.L., Fischer, M.C., Foote, A.D., Franchini, P., Jiggins, C.D., Jones, F.C., Lindholm, A.K., Lucek, K., Maan, M.E., Marques, D.A., Martin, S.H., Matthews, B., Meier, J.I., Möst, M., Nachman, M.W., Nonaka, E., Rennison, D.J., Schwarzer, J., Watson, E.T., Westram, A.M. & Widmer, A. (2014) Genomics and the origin of species. *Nature Reviews Genetics*, **15**, 176–192.

Sick, K. (1961) Haemoglobin polymorphism in fishes. *Nature*, **192**, 894–896.

Sick, K. (1965) Haemoglobin polymorphism of cod in the North Sea and the North Atlantic Ocean. *Hereditas*, **54**, 49–69.

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

Slatkin, M. (2008) Linkage disequilibrium - Understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, **9**, 477–485.

Smadja, C.M. & Butlin, R.K. (2011) A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, **20**, 5123–5140.

Smith, J.M. & Haigh, J. (1974) The hitch-hiking effect of a favourable gene. *Genetical research*, **23**, 23–35.

Sodeland, M., Jorde, P.E., Lien, S., Jentoft, S., Berg, P.R., Grove, H., Kent, M.P., Arnyasi, M., Olsen, E.M. & Knutsen, H. (2016) "Islands of Divergence" in the Atlantic Cod Genome Represent Polymorphic Chromosomal Rearrangements. *Genome biology and evolution*, **8**, 1012–1022.

Solbakken, M.H., Rise, M.L., Jakobsen, K.S. & Jentoft, S. (2016a) Successive losses of central immune genes characterize the Gadiformes' alternate immunity. *Genome biology and evolution*, Early view, evw250.

Solbakken, M.H., Tørresen, O.K., Nederbragt, A.J., Seppola, M., Gregers, T.F., Jakobsen, K.S. & Jentoft, S. (2016b) Evolutionary redesign of the Atlantic cod (*Gadus morhua* L.) Toll-like receptor repertoire by gene losses and expansions. *Scientific reports*, **6**:25211.

Star, B., Nederbragt, A.J., Jentoft, S., Grimholt, U., Malmstrøm, M., Gregers, T.F., Rounge, T.B., Paulsen, J., Solbakken, M.H., Sharma, A., Wetten, O.F., Lanzén, A., Winer, R., Knight, J., Vogel, J.-H., Aken, B., Andersen, O., Lagesen, K., Tooming-Klunderud, A., Edvardsen, R.B., Tina, K.G., Espelund, M., Nepal, C., Previti, C., Karlsen, B.O., Moum, T., Skage, M., Berg, P.R., Gjøen, T., Kuhl, H., Thorsen, J., Malde, K., Reinhardt, R., Du, L., Johansen, S.D., Searle, S., Lien, S., Nilsen, F., Jonassen, I., Omholt, S.W., Stenseth, N.C. & Jakobsen, K.S. (2011) The genome sequence of Atlantic cod reveals a unique immune system. *Nature*, **477**, 207–210.

Stenseth, N.C. & Dunlop, E.S. (2009) Evolution: Unnatural selection. *Nature*, **457**, 803–804.

Storz, J.F. (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

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

Terekhanova, N.V., Logacheva, M.D., Penin, A.A., Neretina, T.V., Barmintseva, A.E., Bazykin, G.A., Kondrashov, A.S. & Mugue, N.S. (2014) Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. (ed CL Peichel). *PLoS Genetics*,

10:e1004696.

Thompson, M.J. & Jiggins, C.D. (2014) Supergenes and their role in evolution. *Heredity*, **113**, 1–8.

Thorsteinsson, V., Pálsson, Ó.K., Tómasson, G.G., Jónsdóttir, I.G. & Pampoulie, C. (2012) Consistency in the behaviour types of the Atlantic cod: repeatability, timing of migration and geo-location. *Marine Ecology Progress Series*, **462**, 251–260.

Tigano, A. & Friesen, V.L. (2016) Genomics of local adaptation with gene flow. *Molecular Ecology*, **25**, 2144–2164.

Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R.S.T., Hecht, J., Knaust, F., Belkhir, K., Klages, S., Dieterich, R., Stueber, K., Piferrer, F., Guinand, B., Bierne, N., Volckaert, F.A.M., Bargelloni, L., Power, D.M., Bonhomme, F., Canário, A.V.M. & Reinhardt, R. (2014) European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature communications*, **5**:5770.

Trucchi, E., Mazzarella, A.B., Gilfillan, G.D., Lorenzo, M.T., Schönswetter, P. & Paun, O. (2016) BsRADseq: screening DNA methylation in natural populations of non-model species. *Molecular Ecology*, **25**, 1697–1713.

Turner, T.L., Hahn, M.W. & Nuzhdin, S.V. (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, **3**:e285.

Twyford, A.D. & Friedman, J. (2015) Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution*, **69**, 1476–1486.

Tørresen, O.K., Star, B., Jentoft, S., Reinar, W.B., Grove, H., Miller, J.R., Walenz, B.P., Knight, J., Ekholm, J.M., Peluso, P., Edvardsen, R.B., Tooming-Klunderud, A., Skage, M., Lien, S., Jakobsen, K.S. & Nederbragt, A.J. (2016) An improved genome assembly uncovers a prolific tandem repeat structure in Atlantic cod. *bioRxiv*:060921.

Via, S. (2001) Sympatric speciation in animals: the ugly duckling grows up. *Trends in ecology & evolution*, **16**, 381–390.

Via, S. (2009) Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America*, **106 Suppl 1**, 9939–9946.

Via, S. (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **367**, 451–460.

Via, S. & West, J. (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.

Via, S., Conte, G., Mason-Foley, C. & Mills, K. (2012) Localizing F_{ST} outliers on a QTL map reveals evidence for large genomic regions of reduced gene exchange during speciation-with-gene-flow. *Molecular Ecology*, **21**, 5546–5560.

Vitalis, R., Dawson, K. & Boursot, P. (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811–1823.

Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y.-C., Shoemaker, D. & Keller, L. (2013) A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*, **493**, 664–668.

Ward, R.D., Woodwark, M. & Skibinski, D.O.F. (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fish Biology*, **44**, 213–232.

Wargelius, A., Leininger, S., Skaftnesmo, K.O., Kleppe, L., Andersson, E., Taranger, G.-L., Schulz, R.W. & Edvardsen, R.B. (2016) Dnd knockout ablates germ cells and demonstrates germ cell independent sex differentiation in Atlantic salmon. *Scientific reports*, **6**:21284.

Wheat, C.W. (2010) Rapidly developing functional genomics in ecological model systems via 454 transcriptome sequencing. *Genetica*, **138**, 433–451.

Williams, G.C. (1975) *Sex and Evolution*. Princeton University Press.

Wolf, J.B.W. & Ellegren, H. (2016) Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*. Early view, Doi: 10.1038/nrg.2016.133

Wright, S. (1921) Systems of Mating. II. the Effects of Inbreeding on the Genetic Composition of a Population. *Genetics*, **6**, 124-143.

Wright, S. (1922) Coefficients of inbreeding and relationship. *The American Naturalist*, **645**, 330–338.

Wu, C.-I. (2001) The genic view of the process of speciation. *Journal of evolutionary biology*, **14**, 851–865.

Yeaman, S. & Whitlock, M.C. (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897–1911.

Zinzow-Kramer, W.M., Horton, B.M., McKee, C.D., Michaud, J.M., Tharp, G.K., Thomas, J.W., Tuttle, E.M., Yi, S. & Maney, D.L. (2015) Genes located in a chromosomal inversion are correlated with territorial song in white-throated sparrows. *Genes, brain, and behavior*, **14**, 641–654.