DTU Library

Genetic analyses reveal complex dynamics within a marine fish management area

Hansen, Jakob Hemmer; Hüssy, Karin; Baktoft, Henrik; Huwer, Bastian; Bekkevold, Dorte; Haslob, Holger; Herrmann, Jens-Peter; Hinrichsen, Hans-Harald; Krumme, Uwe; Mosegaard, Henrik; Eg Nielsen, Einar; Reusch, Thorsten B. H.; Storr-Paulsen, Marie; Velasco, Andres; von Dewitz, Burkhard; Dierking, Jan; Eero, Margit

DR. JAKOB HEMMER-HANSEN (Orcid ID : 0000-0002-1033-2565)

DR. DORTE BEKKEVOLD (Orcid ID : 0000-0002-5297-032X)

PROF. THORSTEN B.H. REUSCH (Orcid ID : 0000-0002-8961-4337)

Genetic analyses reveal complex dynamics within a marine fish management area

Jakob Hemmer-Hansen (1; jhh@aqua.dtu.dk), Karin Hüssy (3; kh@aqua.dtu.dk), Henrik

Baktoft (1; hba@aqua.dtu.dk), Bastian Huwer (3; bhu@aqua.dtu.dk), Dorte Bekkevold (1;

db@aqua.dtu.dk), Holger Haslob (4; holger.haslob@thuenen.de), Jens-Peter Herrmann (4;

jpherrmann@uni-hamburg.de), Hans-Harald Hinrichsen (2; hhinrichsen@geomar.de), Uwe

Krumme (5; uwe.krumme@thuenen.de), Henrik Mosegaard (3; hm@aqua.dtu.dk), Einar Eg

Nielsen (1; een@aqua.dtu.dk), Thorsten B. H. Reusch (2; treusch@geomar.de), Marie Storr-

Paulsen (3; msp@aqua.dtu.dk), Andres Velasco (5; andres.velasco@thuenen.de), Burkhard

von Dewitz (2; bdewitz@geomar.de), Jan Dierking (2; jdierking@geomar.de), Margit Eero

(3; mee@aqua.dtu.dk)

1. Technical University of Denmark, National Institute of Aquatic Resources, 8600 Silkeborg,

Denmark

2. GEOMAR Helmholtz Center for Ocean Research Kiel, Evolutionary Ecology of Marine

Fishes, Düsternbrooker Weg 20, 24105 Kiel, Germany

3. Technical University of Denmark, National Institute of Aquatic Resources, Kemitorvet, Building 202, 2800 Kgs. Lyngby, Denmark

4. Thünen Institute of Sea Fisheries, Palmaille 9, 22767 Hamburg-Altona, Germany

5. Thünen Institute of Baltic Sea Fisheries, Alter Hafen Süd 2, 18069 Rostock, Germany

Corresponding author

Jakob Hemmer-Hansen, jhh@aqua.dtu.dk

Abstract

Genetic data have great potential for improving fisheries management by identifying the fundamental management units – i.e. the biological populations - and their mixing. However, so far the number of practical cases of marine fisheries management using genetics has been limited. Here, we used Atlantic cod in the Baltic Sea to demonstrate the applicability of genetics to a complex management scenario involving mixing of two genetically divergent populations. Specifically, we addressed several assumptions used in the current assessment of the two populations. Through analysis of 483 single nucleotide polymorphisms (SNPs) distributed across the Atlantic cod genome we confirmed that a model of mechanical mixing, rather than hybridization and introgression, best explained the pattern of genetic differentiation. Thus, the fishery is best monitored as a mixed-stock fishery. Next, we developed a targeted panel of 39 SNPs with high statistical power for identifying population of origin and analysed more than 2000 tissue samples collected between 2011 and 2015 as well as 260 otoliths collected in 2003/2004. These data provided high spatial resolution and

allowed us to investigate geographical trends in mixing, to compare patterns for different life stages and to investigate temporal trends in mixing. We found similar geographical trends for the two time points represented by tissue and otolith samples and that a recently implemented geographical management separation of the two populations provided a relatively close match to their distributions. In contrast to the current assumption, we found that patterns of mixing differed between juveniles and adults, a signal likely linked to the different reproductive dynamics of the two populations. Collectively, our data confirm that genetics is an operational tool for complex fisheries management applications. We recommend focussing on developing population assessment models and fisheries management frameworks to capitalize fully on the additional information offered by genetically assisted fisheries monitoring.

Introduction

Genetic and genomic information is increasingly used to inform conservation and management efforts targeting natural populations (Funk *et al.* 2012; Benestan *et al.* 2016). In marine fisheries, management units often do not match the distribution areas of biological populations (Reiss *et al.* 2009, Kerr *et al.* 2017). Increased access to genomic data is a promising development for identifying management units and connectivity patterns in species showing low levels of population structuring (Hemmer-Hansen *et al.* 2014; Gagnaire *et al.* 2015; Kelley *et al.* 2016). Large genomic data sets facilitate the identification of genetic loci

with high discriminatory power (Nielsen *et al.* 2012) which are useful for fast and cost-effective identification of populations and for quantifying their relative contributions to mixed-stock fisheries, i.e. fisheries targeting more than one biological population. These developments facilitate the transfer from proof-of-concept to operational tools which can support and improve management in complex management scenarios. Yet, apart from classical cases such as Pacific salmonids (Dann *et al.* 2013), only few examples of genetically based marine fisheries management exist (Bernatchez *et al.* 2017, Dahle *et al.* 2018).

Atlantic cod (*Gadus morhua*) is a marine fish species of high ecological and commercial importance distributed across the northern Atlantic. It has received considerable interest from studies focusing on basic as well as applied research in both aquaculture and wild populations, and the separation of coastal and offshore populations in the Northern Atlantic (Wennevik *et al.* 2008; Dahle *et al.* 2018) represents an illustrative case of genetically informed marine fisheries management. A number of studies have identified Atlantic cod populations in the brackish water Baltic Sea as genetically unique (Nielsen *et al.* 2001; Berg *et al.* 2015), presumably as a result of a combined colonization and adaptation process to the unique Baltic Sea environment (Nielsen *et al.* 2003; Nielsen *et al.* 2009; Hemmer-Hansen *et al.* 2013), and possibly reflecting separation prior to the opening of the connection between the Atlantic and the Baltic Sea following the last glaciation (Sick 1965). Adaptations to the Baltic Sea environment involve several physiological and life-history traits such as egg buoyancy and spawning time (Nissling *et al.* 1994; Hinrichsen *et al.*, 2012; Petereit *et al.*, 2014). Populations in the environmental transition zone between the North Sea and Baltic Sea have been found to be of mixed genetic ancestry (Nielsen *et al.* 2003; Berg *et al.* 2015), resembling a hybrid zone between pure North Sea and Baltic Sea populations (Nielsen *et al.* 2003). The major genetic break between the Baltic Sea populations is localized in the Arkona Basin region in the western part of the Baltic Sea (see Figure 1 and Nielsen *et*

al. 2003; Poćwierz-Kotus *et al.* 2015). The two cod populations show marked phenotypic differences, with eastern Baltic cod growing more slowly, maturing at a smaller size and attaining lower weight at a given length than western Baltic cod (Berner and Vaske 1985; Bagge *et al.* 1994; Köster *et al.* 2017).

In terms of fisheries management, cod in the Baltic Sea is managed as two stocks, "western" and "eastern" Baltic cod, with the Arkona Basin region managed as part of the western stock until recently (Figure 1). New genetic data as well as otolith morphology and trends in biological characteristics, such as mean weight of cod (Eero *et al.* 2014; Hüssy *et al.* 2016a), have recently demonstrated that mixing of eastern and western (transition zone) cod populations occurs in the Arkona Basin region. Population mixing in this area was of considerable concern due to biased stock assessments, which were conducted separately for management areas east and west of a 13° E border line (see Figure 1). Consequently, mixing was believed to occur within the western management area and, until recently, the assessment of the fisheries resource did not contain detailed information about the spatial and temporal scale of mixing (Eero *et al.* 2014). The marked life-history differences between cod from the two biological populations in the mixing zone could act to decouple population dynamics for the two units with potential consequences for stock assessment and management. For example, differences in spawning time may result in different recruitment patterns in the two populations, and different population productivity may lead to overexploitation of the weaker population component in the mixing zone. As a result, a combination of morphological and genetic data was applied to determine the proportions of eastern and western Baltic cod in fisheries catches taken within the western management area for stock assessment purposes (Hüssy *et al.* 2016a; ICES 2017, Figure 1).

The stock separation used in the current management model was implemented in 2015 (ICES 2015; Hüssy *et al.* 2016a) and has a number of assumptions. First, it assumes that while individuals of the two populations co-occur in the same geographical area (termed "mechanical mixing" in the following), they do not interbreed and hence a possible scenario of hybridization is disregarded. In this geographical region, early studies suggested the presence of mechanical mixing of two populations based on sample statistics of haemoglobin polymorphism data (Sick 1965), while more recent work applying genetic markers with higher statistical power for individual based analyses, but also lower sample sizes, was not able to differentiate between mechanical mixing and hybridization scenarios (Nielsen *et al.* 2003). Although mechanical mixing was recently demonstrated with the use of highly powerful single nucleotide polymorphism genetic markers (Eero *et al.* 2014) the geographical extent of mixing and the possible occurrence of hybrids have so far not been rigorously assessed with tools with high statistical power. Secondly, the current assessment and management model only uses a coarse model of different mixing in two sub-areas (see Figure 1) to account for the geographical variation in mixing proportions. Thirdly, it is currently assumed that mixing proportions are similar for adults and juveniles, and similar proportions of eastern and western fish are allocated to all fish ages when performing the stock assessments. Consequently, the presence of age specific mixing proportions may result in biased stock assessments. Finally, so far there is no mechanistic understanding of the patterns driving spatial patterns and temporal fluctuations in mixing proportions.

Here, we use genetic markers to assess the validity of these assumptions. First, we apply 483 SNP markers in a limited number of individuals and a reduced but specifically selected panel of 39 SNP markers in a large number of individuals to verify the hypothesis of mechanical mixing (i.e. the absence of hybrids) of the two populations. Subsequently, we use the reduced, so-called "high-graded", panel of markers to study population mixing and

dynamics with high spatial and temporal resolution in a large number of individuals collected within the mixing zone. We use these data to investigate if current management assumptions are accurately describing the dynamics of mixing in the region, i.e. addressing the assumptions related to geographical patterns of mixing for juvenile and adults. Finally, we use the genetic data in combination with environmental data to assess possible mechanisms driving population mixing.

Materials and Methods

Sampling and DNA extraction

The geographical focus of this study was at the Arkona Basin region in the Baltic Sea (Figure 1) where previous work suggested that eastern and western Baltic Sea cod populations mix. Baseline samples for the study were represented by Atlantic cod collected at spawning time in the western Baltic Sea and in the Bornholm Basin in the eastern Baltic Sea (Table 1 and Figure 1). These baselines are termed "western" and "eastern" Baltic Sea cod throughout the study. Tissue samples (gills and fins) from 2042 individuals from the mixing zone were collected from research cruises using bottom trawling gear during a 5 year period from 2011 to 2015 (Table 2) and stored in ethanol. For each fish, biological data such as length, weight and maturity stage determined from gonadal maturation status was recorded. Gonadal maturation stage was used to categorize fish as either juvenile or adult and to determine if adult fish were in spawning condition at the time of capture. All fish sampled in November 2013 and February 2014 were below 20 cm length and were therefore assumed to be juveniles. DNA was extracted from tissue samples by Chelex resin (Estoup *et al.* 1996). In addition, we analysed DNA extracted from 260 otoliths collected in 2003 and 2004 to examine temporal stability of population mixing (Table 3). For otoliths, we followed the

procedure outlined in Bonanomi *et al.* (2015), including DNA extraction in a clean laboratory facility and assessment of contamination by microsatellite genotyping prior to SNP genotyping. All samples were genotyped for 39 SNP markers on a Fluidigm Biomark[TM] HD System.

Development of a SNP panel for population assignment

We used a previously published data set composed of more than 1200 SNP loci (see Nielsen *et al.* 2012 for details) to identify loci with high levels of population differentiation between eastern and western Baltic Sea baseline population samples collected in 2007, ranking loci based on estimates of F_{ST} (Weir & Cockerham 1984). We only included SNPs within the same linkage group if the correlation between markers was very low ($r^2<0.01$). In addition, markers on different linkage groups were also only selected if they showed low levels of linkage disequilibrium (max $r^2<0.25$) to ensure that all markers would provide independent information. This procedure also limited biases for population assignment from including linked markers in the well-described highly differentiated genomic regions (at least one corresponding to major chromosomal inversions, Kirubakaran *et al.* 2016) in the cod genome, some of which show high levels of differentiation for the Baltic Sea and transition zone populations (Berg *et al.* 2015). Baseline samples collected in 1996/1997 and 2007 were not significantly different and were pooled as assignment baselines to get a more complete representation of genetic variation in the two populations. Assignment to the most likely baseline (eastern or western Baltic Sea) was based on genotype likelihoods (following Rannala and Mountain 1997), which were used to calculate assignment scores in the programme GeneClass2 (Piry *et al.* 2004). The assignment score is the ratio of the maximum genotype likelihood to the sum of all likelihoods (here the sum of the likelihoods in the two

possible baseline samples). Individuals were assigned to the baseline with the highest score. Statistical power for assignment was evaluated in baseline samples collected in 2012 by self-assignment using the leave-one-out procedure. By estimating power in a different sample ("hold-out sample") than the one used for locus selection, we avoided common problems of "high grading bias" (Anderson 2010). Assignment power was evaluated by estimating the number of misassigned individuals in the baseline samples and by examining the distribution of likelihood ratios, following Ogden and Linacre (2015).

Distinguishing between hybridization and mechanical mixing

To attain high statistical power for distinguishing between hybrids and pure parental genotypes, we re-analysed the Arkona Basin sample collected in 1996 from Nielsen *et al.* (2003) with the use of SNP markers (Table 2; see Nielsen *et al.* 2012 and supplementary file S1 for details of the data set). In the original publication (Nielsen *et al.* 2003) it was not possible to distinguish between the two scenarios with the use of 9 microsatellite markers. Here, we selected SNPs to be distributed throughout the cod genome by selecting markers with a minimum distance of 1 cM on the cod linkage map (Borza *et al.* 2010; Hubert *et al.* 2010). This procedure resulted in a data set consisting of 483 markers. We used the western and eastern Baltic Sea baseline samples collected in 1996/1997 and 2007 as parental populations (Table 1) and used R (R core team 2016) to simulate pure parental and F1 hybrid individuals by drawing alleles with probabilities equal to baseline sample allele frequencies. The simulated individuals were analysed with the Arkona Basin sample in the model based clustering programme STRUCTURE (Hubisz *et al.* 2009) using prior population information for the parental populations. Individual admixture coefficients and their 95% confidence intervals (CIs) were compared for simulated and Arkona Basin individuals. Due to limited

statistical power with a reduced number of markers the same approach could not be used for the majority of individuals in this study as they were only genotyped with 39 SNPs. However, the large number of fish analysed permitted an examination of the distribution of assignment scores which could still be used for estimating the extent of hybridization. Large numbers of fish with intermediate assignment scores would indicate frequent hybridization among the analysed individuals.

Modelling the geographical distribution of mixing

We used binomial generalized linear models (GLM) to examine geographical patterns of mixing for juveniles and adults and the relationships between mixing proportions and potential environmental drivers of mixing. Specifically, we modelled the distribution of mixing proportions while accounting for spatial correlation induced by the sampling design using the following base model extended to include geographical and/or age class (juvenile vs. adult) covariates according to Table 4:

$$\text{East}_{ip} \sim \text{Binomial}(\pi_{ip}, N_{ip})$$

$$E(\text{East}_{ip}) = \pi_{ip} * N_{ip}$$

$$\text{var}(\text{East}_{ip}) = \pi_{ip} * N_{ip} * (1 - \pi_{ip})$$

$$\text{logit}(\pi_{ip}) = \text{salinity}_{ip} + \text{oxygen}_{ip} + \text{temperature}_{ip} + u_{ip}$$

$$u_{ip} \sim \text{GMRF}(0, \textstyle\sum_p)$$

In these models, number of eastern fish, *East$_{ip}$*, out of N_{ip} fish from sampling event (i.e. trawl haul) i at sampling period (defined as year-quarter combinations) p is assumed to follow a binomial distribution with probability π_{ip}. Expected values and variance of *East$_{ip}$* are defined

by $E(East_{ip})$ and $var(East_{ip})$. Salinity, oxygen content and water temperature in the Arkona Basin at the time and location of catch were extracted from the hydrodynamic Kiel Baltic Sea Ice-Ocean Model (BSIOM, Lehmann & Hinrichsen, 2000 and Lehmann *et al.*, 2002; see supplementary file S2 for details) and included as main effects $salinity_{ip}$, $oxygen_{ip}$, $temperature_{ip}$ in the linear predictor function. These continuous covariates were standardized using $x_{std} = (x - mean(x)) / sd(x)$ prior to modelling.

To assess assumptions currently used for management, we fitted five models (Table 4). M0 is the base model as described above and functions as a reference model. Current management assumes that a split into two sub-areas (west and east of 13° E) captures the geographical variation in mixing proportions and that proportions are similar for adults and juveniles. These assumptions are evaluated in models M1 and M2, respectively. M1 includes factor *Area*, describing whether observation *ip* are from the area west or east of 13° E, whereas M2 includes factor *Juvenile*, describing whether observation *ip* pertains to juvenile or adult fish. Due to sparse number of samples containing adults from the western area, the interaction between *Area* and *Juvenile* could not be tested. In M3, a second order random walk model $f(utmX_{ip})$ was included to capture an overall (potentially non-linear) east-west trend on π_{ip} as the study area extension was primarily in the east-west direction. This was extended further in M4 by allowing the second order random walk to vary between juveniles and adults. Models were compared using the Watanabe-Akaike information criteria (WAIC; Zuur *et al.* 2017). The term u_{ip} is a random intercept assumed to be spatially correlated with mean 0 and sampling period specific covariance matrix $\sum_p$, i.e. a replicated spatial random field that is allowed to differ between sampling periods *p*. u_{ip} is assumed to be Markovian and thus follow a Gaussian Markov random field (GMRF) and the associated covariance matrices are modelled using Matérn correlation function and numerically approximated using SPDE (continuous domain stochastic partial differential equation). A logit link function was used to

model π_{ip} as a function of the linear predictor function. The model was fitted in a Bayesian framework using R (R core team 2016) and the package R-INLA (Rue *et al.* 2009; Lindgren *et al.* 2011) following Zuur *et al.* (2017). Diffuse priors were used for all parameters. Model validation and diagnostics were performed using Pearson residuals on population level (i.e. excluding random effects) following procedures as described in Zuur *et al.* (2017). Only data from stations within the mixing zone in Figure 1 were included in the GLM model. Additionally, data sampled prior to 2011 were excluded as the geographic resolution of these was inadequate for this model, and data from two stations were excluded as environmental covariates were unavailable for these. It should be noted that attempts to model a system as complex and dynamic as the mixing of Baltic Sea populations using simple predictors as done here inevitably will be an oversimplification. However, we believe that the present data set does not warrant further complexity in the models considering the available sample sizes.

Our sampling design also allowed a more qualitative assessment of potential effects of major environmental changes observed in January 2003 and December 2014, where strong inflow of high salinity water to the Baltic Sea was observed (Major Baltic Inflow event, Mohrholz *et al.* 2015), by comparing the geographical patterns of mixing before and after inflows.

Results

Development of a SNP panel for population assignment

We identified 39 SNP markers which provided high statistical power for population assignment to eastern and western Baltic cod populations (see supplementary file S3 for detailed information). Because of the different selection criteria used for the two SNP panels, only 23 of the markers in this panel overlapped with markers in the 483 SNP data set used to

differentiate between hybridization and mechanical mixing below (see supplementary files S1 and S3). For the panel of 39 SNPs, only one individual from the western Baltic Sea baseline was assigned to the eastern Baltic Sea baseline, corresponding to a misassignment rate of app. 0.3%. No individuals were misassigned from the eastern Baltic baseline to the western Baltic Sea baseline. Furthermore, the distributions of likelihood ratios were well separated, demonstrating that assignment was unambiguous for the majority of baseline fish (Supplementary file S4). The "average" fish in the western Baltic baseline was more than 22 million times more likely to assign to west than to east, while the "average" fish in the eastern Baltic baseline was more than 2 million times more likely to assign to east than to west.

Distinguishing between hybridization and mechanical mixing

For the 483 SNP data set, the estimation of individual admixture coefficients showed that 95% CIs for simulated parental individuals always overlapped with either 0 or 1 (with 1 corresponding to pure eastern and 0 to pure western origin) but never both. Simulated F1 hybrids had intermediate admixture coefficients and 95% CIs which never included 0 or 1 (Figure 2b,c). The admixture coefficients for the 40 Arkona Basin individuals almost exclusively overlapped with either 0 or 1 (Figure 2a), thus supporting that they represented genotypes with pure parental origins. Only one individual showed an intermediate admixture coefficient with 95% CI not overlapping with 0 or 1. For the 39 SNP data set with 2270 successfully genotyped individuals, only 28 individuals (~1%) had assignment scores below 90 (Supplementary file S5), suggesting the presence of very few individuals of admixed genetic ancestry since a value close to 50 would be expected for F1 hybrids.

Assignment of fish from the mixing zone

Assignment of cod collected between 2011 and 2015 showed mechanical mixing of eastern and western cod in the study region (Figure 3). Analyses of samples from 1996, 2003 and 2004 confirmed the presence of both populations also in these years (Figure 2 and Table 3). Furthermore, there was a geographical gradient from east to west with higher proportions of eastern fish in the collections in the central and eastern parts of the area (Figure 3 and Table 3). In the first quarter of the year, fish in spawning condition represented both western and eastern cod, whereas collections of second and third quarter spawning fish were almost exclusively of eastern origin (Figure 4). Analyses of the origins of juveniles showed that mixing patterns were highly dynamic (Figure 5 for data pooled across years). For example, a difference in geographical patterns was evident for juveniles, where western fish dominated the central parts of the mixing zone in quarters 1-3 while eastern fish were more frequent in the fourth quarter of the year (Figure 5). The size distributions of the juvenile fish from the two populations were also markedly different (Figure 5), showing clear patterns of discrete year-classes of eastern and western fish at the different times of the year. It should be noted that data were pooled across years to increase total sample size, that sample size was limited for the second quarter and that the first quarter was dominated by sampling fish < 20 cm. Thus, it is likely that we have missed larger juveniles in the first quarter of the year. Still, the bimodal distributions for both eastern and western juveniles observed in some of the quarters indicated that collections included at least two year-classes of juveniles from the populations (Figure 5).

Modelling the geographical distribution of mixing

Model comparisons showed the best performance of the model including information about life stage (juvenile vs. adult) and using a gradual geographical transition in mixing proportions (M4; Table 4). This model showed a clear effect of longitude and life stage, as illustrated by the modelled geographical distributions of mixing proportions for the two life stages separately (Figure 6). Although 95% credible intervals were wide, the model suggested that the longitude effect was non-linear (Figure 6), and both juveniles and adults showed major geographical shifts from high proportions of eastern fish in the eastern parts to low proportions in the western parts of the study area. A rapid transition from predominantly western to eastern Baltic cod occurred at a longitude around 13° E for adults (Figure 6a) and around 13.5° E for juveniles (Figure 6b). It should, however, be noted that our sampling coverage for adults in the westernmost parts of the study area was relatively poor. When life stage was not taken into account, the model using the area definitions currently used for stock assessment (M1) performed clearly better than a model without geographical information (M0) and similarly to a model allowing a gradient in mixing proportions (M3). None of the models showed significant effects of the environmental variables (O_2, salinity and temperature) as the 95% credible intervals for parameter estimates included 0 (Supplementary file S6). There was no clear increase in proportions of the high salinity adapted western Baltic cod population immediately after Major Baltic Inflow events in January 2003 and December 2014 (Figure 7 and Table 3).

Discussion

Sustainable fisheries management relies critically on our ability to identify biological populations, and to incorporate information about their spatial distributions and mixing into management (Heath *et al.* 2014; Bernatchez *et al.* 2017). Genetic markers specifically selected to be informative for population identification, so-called "high-graded" marker panels, have been shown to provide efficient traceability tools for marine fisheries management (Nielsen *et al.* 2012, Grewe *et al.* 2015; Montes *et al.* 2017). This approach consists of a two-step procedure; an initial screening of genomes for signatures of population divergence and a subsequent filtering and genotyping of the markers carrying strong signals of divergence. Highly divergent genomic regions may be related to adaptive population divergence in response to environmental differences or linked to cryptic population structure, and may not be observed through the use of a random selection of genetic markers (Gagnaire *et al.* 2015). Consequently, such markers may serve as efficient population tags that can be used to identify the population of origin of individuals. However, since the markers may be located in genomic regions affected by selection, it is generally recommended that baselines should be re-evaluated at regular intervals to assess temporal stability of genetic signatures, in particular in situations under dynamic environmental conditions (Nielsen *et al.* 2012).

The data presented in this paper provide high spatial and temporal resolution of population interaction of Atlantic cod populations in the western Baltic Sea. In particular, the combination of high genome coverage in a limited number of samples and data from a high-graded SNP panel genotyped in thousands of individuals was very useful for investigating complex population dynamics including a direct assessment of assumptions currently used in fisheries management. The high-graded SNP panel was specifically tailored to the study questions and provided very high statistical power for identifying the Atlantic cod populations expected to co-occur in the study region. As such, our study also serves as an example of an

application of a minimum marker panel with maximum statistical power in a marine fisheries management case (Nielsen *et al.* 2012).

Assumption: Mechanical mixing is the main form of interaction between populations

In this study, we have confirmed that eastern and western Baltic cod co-occur in the contact zone between eastern and western Baltic Sea cod populations, and that mechanical mixing and not hybridization is the major form of interaction between these populations. Thus, our results support the first important assumption under the current management scheme, and align with an increasing number of mixed stock scenarios recently unveiled in marine fishes with the use of genetic data, e.g. Atlantic herring (Bekkevold *et al.* 2011, 2015) and cod in the northern Atlantic (Therkildsen *et al.* 2013, Bonanomi *et al.* 2015, Johansen *et al.* 2018). Our results also align with early work inferring mechanical mixing in the Arkona Basin region based on sample departures from Hardy-Weinberg expectations in haemoglobin polymorphism data (Sick 1965). More powerful genetic markers should provide resolution at the individual fish level, however previous work was not able to determine the most likely scenario in this geographical area due to a lack of statistical power with the genetic methods applied at the time and a limited sample size in the contact zone between the two populations (9 microsatellite loci and 59 individuals; Nielsen *et al.* 2003). With 483 SNP markers, we were able to move from inferences based on sample summary statistics to analyses on the individual level and hence to differentiate between pure parental and F1 hybrid genotypes, demonstrating the increased power attained with genomic tools in shallow structure scenarios (Hemmer-Hansen *et al.* 2014). Although our 39 SNP panel did not allow unambiguous differentiation between parental and hybrid genotypes at the individual fish level, we found very few intermediate assignment scores among more than 2000 fish analysed with this

marker set. Thus, these data support the results from the 483 SNP analyses, which had higher statistical power for analysis of individual fish but were conducted on a much smaller number of fish. This raises the question of how this type of mixing can be maintained over the years without resulting in hybridization and a break-down of population structure (Taylor *et al.* 2006). Despite overlapping geographical distributions, western and eastern Baltic cod have temporally distinct spawning times. While western Baltic cod spawning is restricted to a few weeks in early spring, eastern Baltic cod spawn over a prolonged period of time peaking in the summer months (Hüssy 2011; Figure 4). Additionally, egg buoyancy differs between the two populations, where the eggs of the western stock require much higher salinities to remain buoyant than those of the eastern stock (Nissling and Westin, 1997; Petereit *et al.* 2014). The environmental conditions are rarely supporting survival of western cod eggs in the Arkona Basin region, the main limiting factors being low temperature during the spawning time of western cod (Köster *et al.* 2017) and sedimentation due to drift towards the east (Hinrichsen *et al.*, 2012; Petereit *et al.*, 2014). During the 2000s, the environmental conditions for reproduction in the Arkona Basin were generally more favourable for eastern than for western Baltic cod. Finally, it is possible that the substantial genomic differentiation between eastern and western Baltic Sea cod populations (Poćwierz-Kotus *et al.* 2015) are also involving genomic incompatibilities (Bierne *et al.* 2011) which are not directly linked to adaptation to the specific environments inhabited by the two populations but still prevent hybridization between them. These mechanisms may be sufficient to prevent extensive hybridization between the two populations. Mechanical mixing in the contact zone in the Arkona Basin region may seem contradictory to previous work which has suggested that the entire geographic transition zone between the North Sea and Baltic Sea is best described as a hybrid zone between pure North Sea and Baltic Sea populations (Nielsen *et al.* 2003). However, it is possible that occasional hybridization events, for example in the short period in early spring

were both populations spawn in the region (Figure 4) or under certain environmental conditions in this highly environmentally dynamic region (e.g. Mohrholz *et al.* 2015), may generate pulses of gene flow which would result in a hybrid zone signature. In fact, we did observe one individual with an intermediate admixture coefficient and 95% CI non-overlapping with 0 or 1, suggesting that hybridization does occur.

Assumption: Geographical patterns of mixing are accurately described by two sub-areas in the mixing zone

The data confirmed the east-west gradient with respect to the proportion of eastern fish in the contact zone (Hüssy *et al.* 2016a). The spatial modelling of the data showed that the co-occurrence is best explained by a gradient and by taking life stage (juvenile vs. adult) into account (model M4). Without life stage information, both the Area model (M1) and the model describing the gradual geographic change (M3) provided better fits than a model without geographical information (M0). In fact, the Area model performed slightly better, and the geographical trends in the data under the best model suggest that the sub-areas currently used for stock assessment (split at 13° E; ICES 2017) describe the geographical trends in these data quite well, at least for the individuals analysed in this study. However, our data also suggest that the current stock assessment approach does not capture the full geographic complexity of mixing (see next section). It should be noted that we could not evaluate models including temporal variation (within and between years) due to limited data, and temporal variation in mixing may thus be found within the region. Seasonal variation in mixing proportions has been observed for Atlantic herring (Bekkevold *et al.* 2011), and also in Atlantic cod in the current study region (Hüssy *et al.* 2016a), although no consistent pattern of variation was observed between years in the latter study.

Our data from archived otoliths collected in 2003 and 2004 indicate that mixing is not a recent phenomenon. This implies that eastern fish were also present in the area at times with much smaller population sizes of eastern Baltic cod than estimated for late 2000s to early 2010s (Eero *et al.* 2014), as also suggested by morphological data collected from otoliths (Hüssy *et al.* 2016a). Our data further suggested the presence of a strong east-west gradient in mixing proportions in samples from 2003/2004, similarly to the pattern observed in contemporary data collected after 2011, although it should be noted that the samples represent much lower geographical resolution than the contemporary data set. Notably, the time series of samples represented a period including dramatic environmental fluctuation in the form of inflows of high saline water to the Baltic Sea (Major Baltic Inflows, MBIs, Morholz *et al.* 2015). Particularly strong MBIs, leading to more saline conditions extending further eastwards in the Arkona Basin, were observed in January 2003 and December 2014, but were not associated with dramatically higher proportions of the presumably high salinity adapted western Baltic population after the inflow (Table 3 and Figure 7). Indeed, statistical modelling of the full contemporary data set showed that none of the examined environmental parameters within the mixing zone (salinity, temperature and oxygen) were associated with the variation in mixing proportions. This suggests that mixing may be driven by population specific dynamics in the two independent populations, and therefore perhaps environmental conditions outside the geographical region studied here, rather than variation in environmental conditions within the Arkona region (see also discussion on juveniles below).

Assumption: Mixing proportions are similar for juveniles and adults

Our data clearly showed a better fit to models when they included information on life stage. Thus, although the two life stages showed similar overall geographic trends (i.e. an east-west gradient), the results also indicated that the major geographical break in mixing proportions may be different for juveniles and adults, at least for the data included in this study. Consequently, the result indicates that the current approach assuming similar proportions for juveniles and adults does not accurately account for the dynamics of mixing, and that proportions should therefore be estimated for both life stages separately to fully track the dynamics of mixing in the region.

The geographical distributions and size distributions of small juveniles from the two populations provide interesting information about the underlying dynamics potentially driving juvenile mixing proportions, and they indicate important effects of differences in reproductive biology and growth in the two populations (Hüssy 2011). The consequences are directly observed as differences in the length distributions of the fish from the two populations (Figure 5 for data combined across years). In the first quarter, we observed two eastern and one western year-class, most likely corresponding to eastern fish born in the previous year's summer (6 months old) and in the summer the year before (18 months old) and western fish born in the previous years' spring (12 months old). These three year-classes are also seen in the third quarter samples, along with the indication of the young-of-the year western cod (<10 cm long) born in the spring of the same year. Here, the eastern year-classes would be 1 and 2 years old, respectively, suggesting that not all eastern fish have attained sexual maturity (i.e. recruited to the adult population) during their second summer. In the fourth quarter, the young western cod were still sampled, but the dominating western year-class from the third quarter was no longer present, indicating that these fish attain sexual maturity during their second winter at an age of 2. Although we interpret the absence of

juveniles of specific ages as evidence of sexual maturation and recruitment to the adult population it should also be noted that these patterns may be affected by migration to shallow water which are not easily accessed by sampling vessels, and that sampling was heterogeneous across years and that pooled data across years may therefore have been affected by differences in year-class strengths between years. Still, the data indicate short-term intra-annual variation associated with the recruitment dynamics of the two populations and that the proportions of mixing of juveniles vary with differences in year-class strength of the populations having different spawning times. Consequently, seasonal differences in mixing proportions may occur. Although spawning of fish from both populations does seem to occur within the geographical regions studied here (Figure 4), the strengths of the year-classes from the different populations may also be related to conditions in the primary spawning areas for the two populations, which are located outside the mixing zone (Figure 1). Disentangling population specific sexual maturation and recruitment to the adult populations is critical for assessment and management purposes. Thus, a greater focus on population specific dynamics of juveniles vs adults is recommended to provide more accurate estimates of the stock status and management advice in this complex mixing area.

Earlier work has used simulations to evaluate possible consequences of disregarding independent population dynamics of identified genetic units within a management area (Heath *et al.* 2014; Goethel and Berger 2017). Typically, the main concern is overexploitation of weaker components if spatial sub-structuring is disregarded (Reiss *et al.* 2009; Bonanomi *et al.* 2015; Goethel and Berger 2017; Kerr *et al.* 2017). Here, we confirm that the two biological units indeed follow unique growth and maturation trajectories on short-term spatial and temporal scales, resulting in independent recruitment of juveniles to the two adult populations. Importantly, since the size distributions of the two populations overlap, it is not possible to track these cohorts through simply examining size distributions. Previous

work has used otolith based estimation of individual juvenile time of birth to identify the population of origin (Oeberst and Böttcher 1998). However, such analyses are relatively time consuming, and, as we have also shown in our study, spawning time of the two populations may be partly overlapping. Consequently, genetic assignment may be an efficient tool for tracking population specific recruitment patterns in future work.

Conclusions

We have shown here that knowledge of population origin can provide an increased understanding of complex dynamics in a marine fishery on mixed populations. We evaluated the validity of current assumptions used in stock assessment of mixing Baltic cod stocks. Our analyses contribute to refining future assessment and management procedures; in particular, in relation to differentiating mixing patterns for juveniles and adults and to examining temporal patterns of mixing with relatively high geographical resolution. Here, genetics provide a fast and operational tool applicable to all life stages and even historical samples (Bonanomi *et al.* 2015). Currently, otolith based shape analyses are regarded a cheaper alternative for population identification than genetics in the western Baltic Sea, but this method requires continuous ground-truthing against a genetic baseline because otolith shapes are likely affected by complex interactions between environment and genetic background (Hüssy *et al.* 2016b). As recent developments in genotyping technology applications promises major reductions in analysis cost per individual (Campbell *et al.* 2015; Aykanat *et al.* 2016), it is likely that genetics may be a primary tool for future marine fisheries management and monitoring (Bernatchez *et al.* 2017), as currently implemented on a routine basis for Pacific salmonids (Dann *et al.* 2013) and Atlantic cod in the northern Atlantic (Dahle *et al.* 2018). Our study highlights the Baltic Sea cod as a relevant case where genetic tools are already

operational. In addition, genetic analyses could be applied to egg and larval stages to examine spawning dynamics and early life stage mortality of the different populations. These data would be important for estimating the contribution of spawning within the mixing zone to the productivity of the two populations.

Several marine fish species show similar complex patterns of population interaction as the Baltic cod examined in this study (Wennevik *et al.* 2008; Reiss *et al.* 2009; Bekkevold *et al.* 2015; Kerr *et al.* 2017; Saha *et al.* 2017; Dahle *et al.* 2018; Johansen *et al.* 2018). As genomic resolution continues to increase in species of interest to fisheries management, it is likely that high powered genetic tools will soon be operational for a range of species. To capitalize fully on this promising development, concomitant advances in stock assessment methods and management frameworks that take complex mixing dynamics into account will be needed.

(FishPopTrace) supported the project. We are grateful for the assistance for sample collection from crews on multiple Danish and German research and monitoring cruises.

Data archiving statement

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.83765sm

Figure legends

Figure 1

Map of the study region showing the geographical location of eastern (blue) and western (red) Baltic Sea population spawning areas used as baselines. The mixing zone is marked by a dashed black line and a dashed yellow line marks the separation of two areas currently used for stock assessment within the mixing zone. Until 2015, the mixing zone belonged exclusively to the western Baltic cod population for stock assessment and management.

Figure 2

Individual admixture proportions with 95% confidence intervals in the sample from the Arkona Basin analysed with 483 SNPs in a), simulated pure parental individuals in b) and simulated F1 hybrids in c). Pure eastern Baltic fish have an admixture coefficient close to 1. Data in b) were simulated assuming a 50/50 mixing ratio.

Figure 3

Proportions of eastern and western cod for all juvenile and adult samples collected from 2011 to 2015.

Figure 4

Proportion of eastern and western cod among spawning fish collected in quarter 1 (n=21) in a), quarter 2 (n=230) in b) and quarter 3 (n=18) in c) in samples collected from 2011 to 2015.

Figure 5

Spatial distributions of mixing proportions and size distributions of juvenile cod for quarters 1 (a and b, n=166), 2 (c and d, n=66), 3 (e and f, n=195) and 4 (g and h, n=200). Overlapping size distributions indicated by shading of colours. Data were combined across years.

Figure 6

Modelled geographical distribution of mixing proportions for adults (a) and juveniles (b) for model M4 conditional on oxygen, salinity and temperature being fixed at mean values (see supplementary file S6). Shaded area describes 95% credible intervals and the geographical (longitudinal) location of sampling stations is shown above the x axis. Note the low geographical coverage for adults in the western part of the study region, which prevented a detailed modelling of the effects of Area and Juvenile.

Figure 7

Proportions of eastern and western cod before (n=229) in a) and after (n=236) in b) a Major

Baltic Inflow of high salinity water in December 2014.

Supporting information

Supplementary file S1.xlsx

Locus names, genomic locations and divergences (F_{ST}) between eastern and western Baltic

Sea cod populations for 483 SNP loci used for investigating hybridization through model

based clustering

Supplementary file S2.pdf

Detailed information about environmental parameters included in generalized linear

modelling

Supplementary file S3.xlsx

Locus names, genomic locations and divergences (F_{ST}) between eastern and western Baltic

Sea cod populations for 39 SNP loci use for population assignment

Supplementary file S4.pdf

Distributions of Log(Likelihood ratios) in 150 eastern and 150 western baseline samples

collected in 2011/2012 based on the 39 SNP assignment panel

Supplementary file S5.pdf

Distribution of assignment scores for individuals analysed with 39 SNP markers. Note that the y axis is log scaled. We added one observation to each assignment score bin to facilitate plotting of observations on a log scale.

Supplementary file S6.pdf

Posterior means and 95 % credible intervals of estimated fixed effects parameters oxygen, salinity and temperature for the five evaluated models.

References

Anderson, E. C. (2010) Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. *Molecular Ecology Resources*, 10, 701-710.

Aykanat, T., Lindqvist, M., Pritchard, V. L., & Primmer, C. R. (2016) From population genomics to conservation and management: a workflow for targeted analysis of markers identified using genome-wide approaches in Atlantic salmon *Salmo salar*. *Journal of Fish Biology*, 89 (SI), 2658-2679.

Bagge, O., Thurow, F., Steffensen, E., & Bay, J. (1994) The Baltic cod. *Dana*, 10, 1–28.

Bekkevold, D., Clausen, L. A. W., Mariani, S., Andre, C., Hatfield, E. M. C., Torstensen, E., Ryman, N., Carvalho, G. R., & Ruzzante, D. E. (2011) Genetic mixed-stock analysis of Atlantic herring populations in a mixed feeding area. *Marine Ecology – Progress Series*, 442, 187-199.

Bekkevold, D., Helyar, S. J., Limborg, M. T., Nielsen, E. E., Hemmer-Hansen, J., Clausen, L. A. W., FishPopTrace Consortium, & Carvalho, G. R. (2015) Gene-associated markers can assign origin in a weakly structured fish, Atlantic herring. *ICES Journal of Marine Science*, 72, 1790-1801.

Benestan, L. M., Ferchaud, A. L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., Schwartz, M. K., Kelley, J. L., & Luikart, G. (2016) Conservation genomics of natural and managed populations: building a conceptual and practical framework. *Molecular Ecology*, 25, 2967-2977.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., Martinsohn, J. T., Miller, K. M., Naish, K. A., Ovenden, J. R., Primmer, C. R., Suk, H. Y., Therkildsen, N. O., & Withler, R. E. (2017) Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in Ecology & Evolution*, 32, 665-680.

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.

Berner, M., & Vaske, B. (1985) Morphometric and meristic characters of cod stocks in the Baltic Sea. ICES Document CM 1985/J: 11.

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

Bonanomi, S., Pellissier, L., Therkildsen, N. O., Hedeholm, R. B., Retzel, A., Meldrup, D., Olsen, S. M., Nielsen, A., Pampoulie, C., Hemmer-Hansen, J., Wisz, M., Grønkjær, P., &

Nielsen, E. E. (2015) Archived DNA reveals fisheries and climate induced collapse of a major fishery. *Scientific Reports*, 5, 15395.

Borza, T., Higgins, B., Simpson, G., & Bowman, S. (2010) Integrating markers Pan I and haemoglobin with the genetic linkage map of Atlantic cod (*Gadus morhua*). *BMC Research Notes*, 3, 261.

Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015) Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, 15, 855-867.

Dahle, G., Johansen, T., Westgaard, J. I., Aglen, A., & Glover, K. A. (2018) Genetic management of mixed-stock fisheries "real-time": The case of the largest remaining cod fishery operating in the Atlantic in 2007-2017. *Fisheries Research*, 205, 77-85.

Dann, T. H., Habicht, C., Baker, T. T., & Seeb, J. E. (2013) Exploiting genetic diversity to balance conservation and harvest of migratory salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 70, 785–793.

Eero, M., Hemmer-Hansen, J., & Hüssy, K. (2014) Implications of stock recovery for a neighboring management unit: experience from the Baltic cod. *ICES Journal of Marine Science*, 71, 1458-1466.

Estoup, A., Largiader, C. R., Perrot, E., & Chourrout, D. (1996) Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology*, 5, 295–298.

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

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

Gagnaire, P. A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., & Bierne, N. (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, 8, 769-786.

Goethel, D. M., & Berger, A. M. (2017) Accounting for spatial complexities in the calculation of biological reference points: effects of misdiagnosing population structure for stock status indicators. *Canadian Journal of Fisheries and Aquatic Sciences*, 74, 1878–1894.

Grewe, P. M., Feutry, P., Hill, P. L., Gunasekera, R. M., Schaefer, K. M., Itano, D. G., Fuller, D. W., Foster, S. D., & Davies, C. R. (2015) Evidence of discrete yellowfin tuna (*Thunnus albacares*) populations demands rethink of management for this globally important resource. *Scientific Reports*, 5, 16916.

Heath, M. R., Culling, M. A., Crozier, W. W., Fox, C. J., Gurney, W. S. C., Hutchinson, W. F., Nielsen, E. E., O'Sullivan, M., Preedy, K. F., Righton, D. A., Speirs, D. C., Taylor, M. I., Wright, P. J., & Carvalho, G. R. (2014) Combination of genetics and spatial modelling highlights the sensitivity of cod (*Gadus morhua*) population diversity in the North Sea to distributions of fishing. *ICES Journal of Marine Science*, 71, 794-807.

Hemmer-Hansen, J., Nielsen, E. E., Therkildsen, N. O., Taylor, M. I., Ogden, R., Geffen, A., 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.

Hemmer-Hansen, J., Therkildsen, N. O., & Pujolar, J. M. (2014) Population genomics of marine fishes: next generation prospects and challenges. *Biological Bulletin*, 227, 117-132.

Hinrichsen, H.-H., Hüssy, K., & Huwer, B., (2012) Spatio-temporal variability in western Baltic cod early life stage survival mediated by egg buoyancy, hydrography and hydrodynamics. *ICES Journal of Marine Science*, 69, 1744-1752.

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.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322-1332.

Hüssy, K. (2011) Review of western Baltic cod (*Gadus morhua*) recruitment dynamics. *ICES Journal of Marine Science*, 68, 1459-1471.

Hüssy, K., Hinrichsen, H.-H., Eero, M., Mosegaard, H., Hemmer-Hansen, J., Lehmann, A., & Lundgaard, L. S. (2016a) Spatio-temporal trends in stock mixing of eastern and western Baltic cod in the Arkona Basin and the implications for recruitment. *ICES Journal of Marine Science*, 73, 293–303.

Hüssy, K., Mosegaard, H., Albertsen, C. M., Nielsen, E. E., Hemmer-Hansen, J., & Eero, M. (2016b) Evaluation of otolith shape as a tool for stock discrimination in marine fishes using Baltic Sea cod as a case study. *Fisheries Research*, 174, 210-218.

ICES. 2015. Report of the Benchmark Workshop on Baltic Cod Stocks (WKBALTCOD), 2–6 March 2015, Rostock, Germany. ICES CM 2015/ACOM: 35. 172 pp.

ICES (2017) Report of the Baltic Fisheries Assessment Working Group (WGBFAS),19-26 April 017, Copenhagen, Denmark. ICES CM 2017/ACOM:11. 810 pp.

Johansen, T., Westgaard, J. I., Seliussen, B. B., Nedreaas, K., Dahle, G., Glover, K. A., Kvalsund, R., & Aglen, A. (2018) "Real-time" genetic monitoring of a commercial fishery on

the doorstep of an MPA reveals unique insights into the interaction between coastal and migratory forms of the Atlantic cod. *ICES Journal of Marine Science*, 75, 1093-1104.

Kelley, J. L., Brown, A. P., Therkildsen, N. O., & Foote A. D. (2016) The life aquatic: advances in marine vertebrate genomics. *Nature Reviews Genetics*, 17, 523-534.

Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L. W., Dickey-Collas, M., Goethel, D. R., Hatfield, E. M. C., Kritzer, J. P., & Nash, R. D. M. (2017) Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES Journal of Marine Science*, 74, 1708-1722.

Kirubakaran, T. G., Grove, H., Kent, M. P., Sandve, S. R., Baranski, M., Nome, T., De Rosa, M. C., Righino, B., Johansen, T., Ottera, 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.

Köster, F. W., Huwer, B., Hinrichsen, H.-H., Neumann, V., Makarchouk, A., Eero, M., Dewitz, B. V., Hüssy, K., Tomkiewicz, J., Margonski, P., Temming, A., Hermann, J. P., Oesterwind, D., Dierking, J., Kotterba, P., & Plikshs, M. (2017) Eastern Baltic cod recruitment revisited-dynamics and impacting factors. *ICES Journal of Marine Science*, 74, 3-19.

Lehmann, A., & Hinrichsen, H.-H. (2000) On the thermohaline variability of the Baltic Sea. *Journal of Marine Systems*, 25, 333-357.

Lehmann, A., Krauß, W., Hinrichsen, H.-H. (2002) Effects of remote and local atmospheric forcing on circulation and upwelling in the Baltic Sea. *Tellus*, 54A, 299–316.

Lindgren, F., Rue, H., & Lindstrom, J. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach (with discussion). *Journal of the Royal Statistical Society, Series B*, 73, 423–498.

Mohrholz, V., Naumann, M., Nausch, G., Krüger, S., & Gräwe, U. (2015) Fresh oxygen for the Baltic Sea - An exceptional saline inflow after a decade of stagnation. *Journal of Marine Systems*, 148, 152–166.

Montes, I., Laconcha, U., Iriondo, M., Manzano, C., Arrizabalaga, H., & Estonba, A. (2017) Reduced Single Nucleotide Polymorphism Panels for Assigning Atlantic Albacore and Bay of Biscay Anchovy Individuals to Their Geographic Origin: Toward Sustainable Fishery Management. *Journal of Agricultural and Food Chemistry*, 65, 4351-4358.

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

Nielsen, E. E., Cariani, A., Mac Aoidh, E., Maes, G. E., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J. *et al.* (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications*, 3, 851.

Nielsen, E. E., Hansen, M. M., Ruzzante, D. E., Meldrup, D., & Grønkjær, P. (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. *Molecular Ecology*, 12, 1497-1508.

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. (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, 9, 276.

Nissling, A., Kryvi, H., & Vallin, L. (1994) Variation in egg buoyancy of Baltic cod *Gadus morhua* and its implications for egg survival in prevailing conditions in the Baltic Sea. *Marine Ecology - Progress Series*, 110, 67-74.

Nissling, A., & Westin, L. (1997) Salinity requirements for successful spawning of Baltic and Belt Sea cod and the potential for cod stock interactions in the Baltic Sea. *Marine Ecology - Progress Series*, 152, 261–271.

Oeberst, R,, & Böttcher, U. (1998) Development of juvenile Baltic cod described with meristic, morphometric and sagitta otolith parameters. International Council for the Exploration of the Sea, ICES C.M. 1998/CC:15.

Ogden, R., & Linacre, A. (2015) Wildlife forensic science: A review of genetic geographic origin assignment. *Forensic Science International: Genetics*, 18, 152–159.

Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L., & Estoup, A. (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity*, 95, 536-539.

Petereit, C., Hinrichsen, H.-H., Franke, A., & Köster, F. W. (2014) Floating along buoyancy levels: Dispersal and survival in western Baltic fish eggs. *Progress in Oceanography*, 122, 131-152.

Pocwierz-Kotus, A., Kijewska, A., Petereit, C., Bernas, R., Wiecaszek, B., Arnyasi, M., Lien, S., Kent, M. P., & Wenne, R. (2015). Genetic differentiation of brackish water populations of cod *Gadus morhua* in the southern Baltic, inferred from genotyping using SNP-arrays. *Marine Genomics*, 19, 17-22.

R Core Team (2016) R: A Language and Environment for Statistical. R Foundation for Statistical Computing, Vienna, Austria.

Rannala, B., & Mountain, J. L. (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the USA*, 94, 9197-9221.

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.

Rue, H., Martino, S., & Chopin, N. (2009) Approximate Bayesian inference for latent Gaussian models using inte-grated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*, 71, 319-392.

Saha, A., Johansen, T., Hedeholm, R., Nielsen, E. E., Westgaard, J. I., Hauser, L., Planque, B., Cadrin, S. X., & Boje, J. (2017) Geographic extent of introgression in *Sebastes mentella* and its effect on genetic population structure. *Evolutionary Applications*, 10, 77-90.

Sick, K. (1965) Haemoglobin polymorphism of cod in the Baltic and the Danish belt sea. *Hereditas*, 54, 19-48.

Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D., & Gow, J. L. (2006) Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, 15, 343-355.

Therkildsen, N. O., Hemmer-Hansen, J., Hedeholm, R. B., Wisz, M. S., Pampoulie, C., Meldrup, D., Bonanomi, S., Retzel, A., Olsen, S. M., & Nielsen, E. E. (2013) Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evolutionary Applications*, 6, 690–705.

Wennevik, V., Jørstad, K. E., Dahle, G., & Fevolden, S. E. (2008) Mixed stock analysis and the power of different classes of molecular markers in discriminating coastal and oceanic

Atlantic cod (*Gadus morhua* L.) on the Lofoten spawning grounds, Northern Norway. *Hydrobiologia*, 606, 7-25.

Weir, B., & Cockerham, C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358–1370.

Zuur, A. F., Ieno, E. N., & Saveliev, A. A. (2017). Beginner's Guide to Spatial, Temporal and Spatial-Temporal Ecological Data Analysis with R-INLA. Volume I: Using GLM and GLMM. Highland Statistics Ltd., Newburgh, United Kingdom.

Table 1. Baseline samples used for assignment and for estimating assignment power

Baseline	Sampling time	Sample size	No SNPs	Use	Source
Eastern	April, 1997	40	483/39	Assignment	Nielsen *et al.* (2012)
Eastern	February, 2007	40	483/39	SNP selection/Assignment	Nielsen *et al.* (2012)
Eastern	July/August, 2011/2012	150	39	Power estimation	This study
Western	February/March, 1996	40	483/39	Assignment	Nielsen *et al.* (2012)
Western	March, 2007	37	483/39	SNP selection/Assignment	Nielsen *et al.* (2012)
Western	February, 2012	150	39	Power estimation	This study

Table 2. Tissue samples analysed from the mixing zone

Year	Month	Quarter	Total sample size	Number of spawning fish	Number of juveniles[1]	Comments
1996	February/March	1	40	0	0	Sample from Nielsen *et al.* (2003) re-analysed with 483 SNPs in this study
2011	June	2	536	152	50	
2013	July	3	41	0	41	
2013	November	4	150	0	150	All fish below 20 cm length
2014	February	1	150	0	150	All fish below 20 cm length
2014	February	1	21	19	0	Spawning fish, not used for environmental correlation
2014	April	2	289	78	19	
2014	August	3	145	0	94	
2014	October	4	229	0	51	
2015	February	1	236	2	16	
2015	July	3	90	18	23	
2015	September	3	155	0	43	

[1]Juveniles were identified through maturation estimation by fish dissection and by assuming that fish smaller than 20 cm were all juveniles.

Table 3. Otolith samples analysed from the mixing zone. Proportions of eastern fish are shown in brackets. A Major Baltic Inflow was observed in January 2003.

Year	Quarter	Western Arkona	Eastern Arkona	Total
2003	1	65 (0.28)		
	4	20 (0.25)	43 (0.93)[1]	
	Total 2003	85	43	128
2004	1	62 (0.16)		
	4	22 (0.09)	48 (0.83)	
	Total 2004	84	48	132

1: Thirty-two of the individuals were collected on a fishing trip to the westernmost part of the neighbouring area, i.e. to the east of the eastern border of the Arkona mixing zone. Among the 11 individuals collected within the Arkona region, the proportion of eastern fish was 0.82.
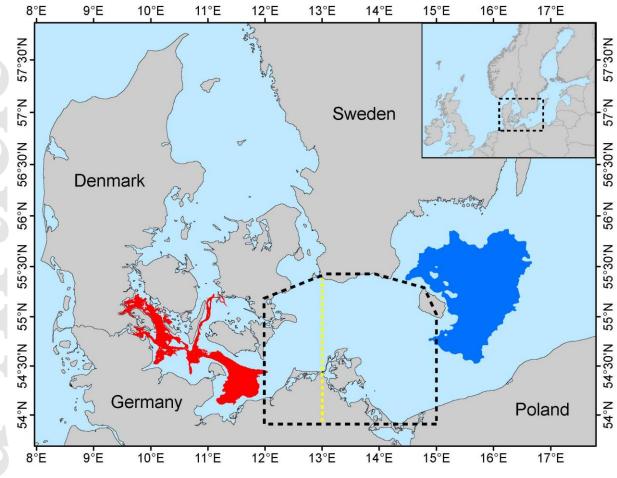
Table 4. Binomial generalized linear models compared for fits to observed mixing proportions in the Arkona Basin region.

Model	Main effects	Geographic and age-class covariates*	WAIC	ΔWAIC#
M0	O2+sal+temp	(none)	620.27	98.28
M1	O2+sal+temp	Area	614.65	92.66
M2	O2+sal+temp	Juvenile	549.21	27.22
M3	O2+sal+temp	$F(utmX)$	615.81	93.8
M4	O2+sal+temp	$F_{juvenile}(utmX)$	521.99	-

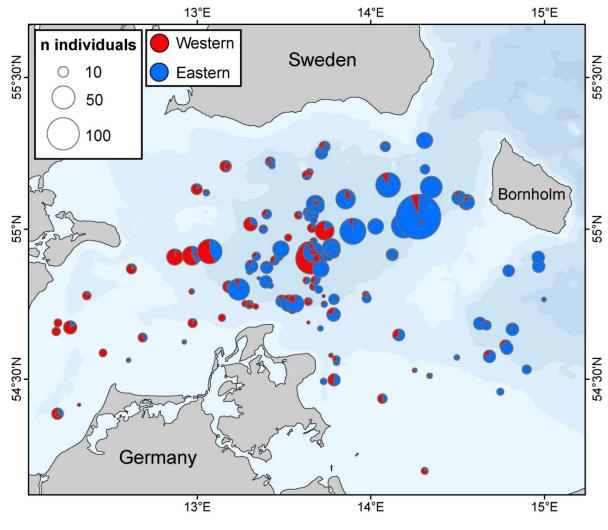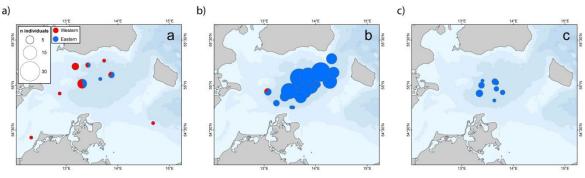* "Area" indicates the area difinitions currently used for stock assessment (see Figure 1). "F(utmX)" is the model including a longitudinal smoother. "Juvenile" indicates models taking variation between juveniles and adults into account.

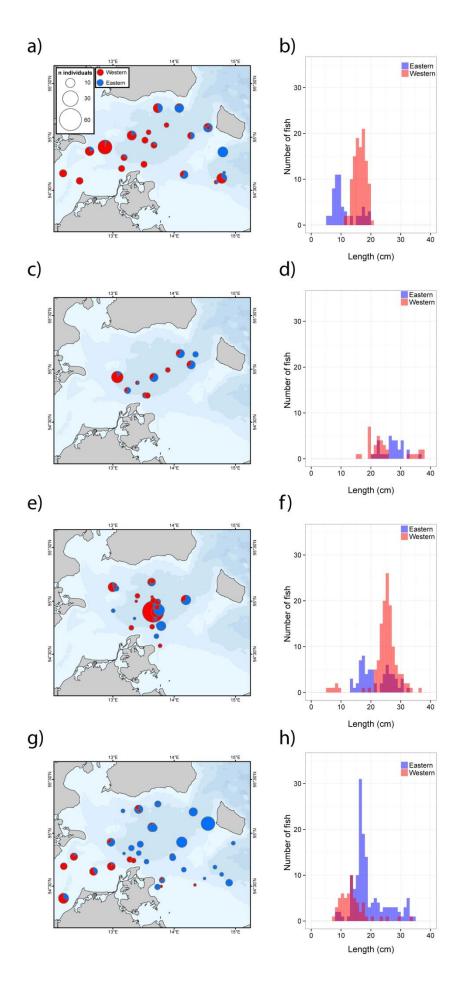# Difference between best model (M4) and current model.

a)

b)

c)

q (eastern Baltic)