ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2009.00779.x

DISENTANGLING THE EFFECTS OF EVOLUTIONARY, DEMOGRAPHIC, AND ENVIRONMENTAL FACTORS INFLUENCING GENETIC STRUCTURE OF NATURAL POPULATIONS: ATLANTIC HERRING AS A CASE STUDY

Oscar E. Gaggiotti,[1] Dorte Bekkevold,[2,3,4] Hanne B. H. Jørgensen,[4,5] Matthieu Foll,[6] Gary R. Carvalho,[7] Carl Andre,[8] and Daniel E. Ruzzante[9]

[1]*Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier, 38041 Grenoble, France*

[2]*National Institute of Aquatic Resources, Technical University of Denmark, Vejlsøvej 39, 8600 Silkeborg, Denmark*

[3]*E-mail: db@aqua.dtu.dk*

[5]*University of Aarhus, Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Blichers Allé 20, PO Box 50, DK-8830 Tjele, Denmark*

[6]*Computational and molecular population genetics lab, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland*

[7]*School of Biological Sciences, Bangor University, Environment Centre Wales, Bangor, Gwynedd LL57 2UW, United Kingdom*

[8]*Department of Marine Ecology-Tjärnö, University of Gothenburg, S-452 96 Stromstad, Sweden*

[9]*Department of Biology, Dalhousie University, Halifax NS B3H 4J1, Canada*

Received February 12, 2009
Accepted June 23, 2009

The spatial structuring of intraspecific genetic diversity is the result of random genetic drift, natural selection, migration, mutation, and their interaction with historical processes. The contribution of each has been typically difficult to estimate, but recent advances in statistical genetics have provided valuable new investigative tools to tackle such complexity. Using a combination of such methods, we examined the roles of environment (i.e., natural selection), random genetic processes (i.e., drift), and demography and life histories (e.g., feeding migrations) on population structure of a widely distributed and abundant marine pelagic fish of economic importance, Atlantic herring (*Clupea harengus*). Individuals were collected during peak spawning time from 19 spawning locations spanning the region from the western North Sea to the eastern Baltic Sea ($N = 1859$, eight microsatellite loci). We carried out separate analyses of neutral and selected genetic variation, which allowed us to establish that the two most important factors affecting population structure were selection due to salinity at spawning sites and feeding migrations. The genetic signal left by the demographic history of herring, on the other hand, seems to have been largely eroded, which is not surprising given the large reproductive potential and presumed enormous local effective population sizes of pelagic fish that constrain the effect of

[4]Authors contributed equally.

stochastic processes. The approach we used can in principle be applied to any abundant and widely distributed aquatic or terrestrial species.

KEY WORDS: Bayesian methods, demographic history, migration, pelagic fish, selection.

A fundamental problem in evolutionary biology is to identify the processes shaping the geographic distribution of intraspecific genetic diversity. Random genetic drift, natural selection, migration and mutation, and their interaction with historical processes are all known to affect the distribution of diversity in space and time. The contribution of these processes to genetic structure has, however, been typically difficult to estimate. Nonetheless, understanding their roles is of paramount importance for several fields of research, including conservation biology, agricultural crop and breed improvement, and identification of disease genes in humans and other economically important species.

Recent major advances in population genomics have fuelled the development of statistical genetic methods that provide the necessary tools for addressing such an objective. For instance, recently developed outlier detection methods (e.g., Beaumont and Nichols 1996; Beaumont and Balding 2004; Foll and Gaggiotti 2008) can be used to discriminate between selected and neutral genetic diversity. The identification of markers under selection is a necessary first step before separating the effects of selection from other evolutionary or demographic processes. A logical follow-up involves the estimation of the relative importance of different environmental factors, life history, and demographic processes. Such is now feasible using a method that combines genetic and nongenetic data for the identification of environmental factors underlying selection and effecting the structuring of genetic variation (Foll and Gaggiotti 2006).

The complexities inherent to the examination of the relative roles of selection, random evolutionary processes, demography and life histories as well as landscape evolution are pervasive among widely distributed and abundant marine species. Geological history, climate, demographic trajectories as well as natural selection have all been invoked as potentially affecting their patterns of genetic diversity (e.g., Palumbi 1994). Additional complexity may be added by anthropogenic impact that may drive significant short-term change (Turner et al. 2006, reviews in Hauser and Carvalho 2008 and in Palstra and Ruzzante 2008). In the present article, we use the statistical genetics methods identified above to examine the roles of environment (i.e., natural selection), random genetic processes (i.e., drift), and demography and life histories (e.g., feeding migrations) on the population structure of a widely distributed and abundant marine pelagic fish of economic importance, Atlantic herring (*Clupea harengus*). We use data from herring spawning aggregations located along a steep environmental gradient from the North Sea to the Inner Baltic Sea.

Atlantic herring is widely distributed in the North Atlantic and genetic population structure has been demonstrated at regional levels in both the eastern and western North Atlantic. Transoceanic differentiation has been shown to be comparable to that within regions (McPherson et al. 2004; Mariani et al. 2005), suggesting genetic divergence cannot be explained simply by isolation-by-distance mechanisms. In the region of our study (North Sea to Baltic Sea), some populations exhibit annual long-distance migrations from their individual spawning grounds to common feeding grounds, leading to aggregations of mixed origin in feeding grounds (e.g., Ruzzante et al. 2006), and potentially facilitating gene flow among spawning populations. Despite these seasonal migrations, however, herring in the North and Baltic Seas appear genetically structured (Bekkevold et al. 2005; Jørgensen et al. 2005), supporting the assumption of natal homing in this species (e.g., Iles and Sinclair 1982; Ruzzante et al. 2006). Natural selection has also been recently implicated as an important factor affecting structure in this species (Larsson et al. 2007). The transition zone between the North Sea and the Baltic Sea is an area with a strong salinity gradient spanning from 30‰ to 34‰ in the North Sea and from 6‰ to 8‰ in the Baltic proper, with levels decreasing down to 3‰ in the innermost (northeastern) parts of the Baltic (Fig. 1). The transition zone also marks a change in water temperature regime from fairly stable in the North Sea to seasonally varying in the Baltic Sea (HELCOM 1996). Recent molecular studies show that this area coincides with steep increases in genetic differentiation in several marine organisms including Atlantic herring (Bekkevold et al. 2005), cod (*Gadus morhua*), and turbot (*Scophthalmus maximus* L. (Nielsen et al. 2003, 2004; respectively), as well as in other taxa (review in Johannesson and André 2006). Bekkevold et al. (2005) reported a correlation between salinity at spawning locations and genetic differentiation among populations from the North Sea to the Western Baltic Sea. Such a correlation between genetic and environmental differentiation was hypothesized to be the result of selection acting through varying environmental conditions on what may historically have been a single population (i.e., primary intergradation) or the result of two ancestral and divergent populations (one in the Baltic Sea and one in the North Sea) that became connected through a zone of secondary contact.

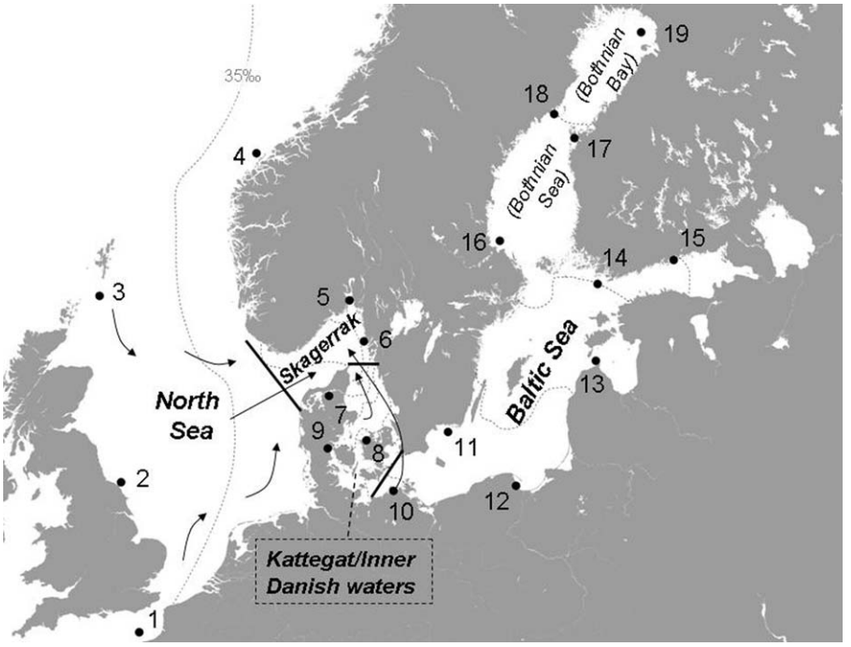

Figure 1. **Sampling locations for *Clupea harengus* (see Table 1 for sample details). The North Sea-Baltic Sea salinity gradient is indicated by stippled isoclines. Arrows indicate main migratory routes to the Skagerrak mixed-stock feeding area for populations from Inner Danish waters (represented by samples 7–9), Rügen (sample 10), and the North Sea (represented by samples 1–3; see text).**

The studies on herring genetics described above suggest that the observed genetic structure of Atlantic herring in the North and Baltic Seas is the result of interplay between environmental factors, behavioral/physiological traits, demographic processes, and geological history, yet no previous effort has attempted to evaluate their relative importance on observed population structure. The purpose here is to address this gap by focusing on three main issues. First, we investigated if natural selection played a role in the genetic structure of Atlantic herring in the North and Baltic Seas. Our results indicate that this is indeed the case. Second, we examined which environmental factor, either salinity or temperature, or their combination is correlated with such an effect. Our results indicate that salinity may influence the structuring of one locus but that temperature plays no detectable role. The population differentiation observed among herring aggregations in the North and Baltic Seas does indeed appear to be related to differences in salinity among spawning aggregations. Third, we investigated the effect of migratory behavior of the different aggregations and their demographic history. Results indicated that population-specific migratory behaviors had a significant effect on structure. For demographic history, we considered two alternative scenarios: (1) a progressive colonization from the North Sea into the Baltic Sea, and (2) a secondary contact zone with two ancestral populations, one in the North Sea and another in the Baltic Sea that progressively colonize the transition zone. Our results support the latter scenario.

Materials and Method

SAMPLES

Data included a total of 19 populations sampled on spawning grounds during peak spawning times, which vary among local components (Table 1, Fig. 1). Temperature and salinity parameters were estimated from average values on the spawning grounds during the month of peak spawning activity (Table 1). Location-specific differences in these values were strongly positively correlated across the spawning, egg, and first larval phases (Bekkevold et al. 2005) which are sensitive to changes in these parameters. Most herring populations undertake annual migrations between spawning and feeding areas, as listed in Table 1 for each population sample. Additional information on samples and environmental data is given in Bekkevold et al. (2005), Jørgensen et al. (2005), and Mariani et al. (2005).

GENETIC DATA

For the present study, we reanalyzed genetic data reported in Bekkevold et al. (2005), Jørgensen et al. (2005), and Mariani et al. (2005), compiling information for eight tetranucleotide microsatellite loci that were common to all three studies. PCR primers for these loci were developed for *C. harengus* by McPherson et al. (2001; Cha1017, Cha1020, Cha1027, Cha1202) and *C. pallasi* by Olsen et al. (2002; Cpa101, Cpa111, Cpa112, Cpa114). Scoring and data quality control measures are presented in Bekkevold et al. (2005) and Jørgensen et al. (2005). Briefly, a

Table 1. *Clupea harengus* sampling locations with spawning time, sample size, average temperature, and salinity spawning site in month of spawning. Feeding areas of populations feeding outside their natal region (see text) are also given. Sample no. refers to Figure 1. Further sample details are found in Mariani et al. (2005) for samples 1–4; in Bekkevold et al. (2005) for 5–10, and in Jørgensen et al. (2005) for 11–19.

Sea/Region	Sampling location (abbreviated)	Sample no.	Lat./ Long.	Sampling/ spawning month	Sampling year	Sample size	Temp. (°C)[1]	Salinity[1]	Feeding area[2]
North Sea	English Channel (CHAN)	1	50.15N/0.43E	Nov.	2003	63	13.0	35.0	Skagerrak (Juv)
	Flamborough Head (FLAM)	2	54.57N/0.45W	Sep.	2003	77	13.3	34.0	Skagerrak (Juv)
	Shetland (SHET)	3	59.46N/1.48W	Aug.	2003	85	11.3	35.2	Skagerrak (Juv)
	Möre (MORE)	4	62.78N/6.08E	Feb.	2003	78	6.5	34.6	Natal region
Skagerrak	Tjöme (TJO)	5	59.35N/10.55E	Mar.	2003	116	3.7	31.2	Natal region
	Flatbrotten (FLAT)	6	58.10N/11.33E	Mar.	2003	100	3.2	30.8	Natal region
Inner Danish waters	Limfjord (LIM)	7	57.06N/10.06E	May	2003	100	6.7	16.0	Skagerrak (Ad)
(IDW)	Kattegat (KAT)	8	55.95N/11.61E	Apr.	2003	100	6.5	21.9	Skagerrak (Ad)
	Lillebælt (LIL)	9	55.66N/9.90E	Apr.	2003	100	5.8	18.0	Skagerrak (Ad)
Western Baltic Sea	Rügen (RUG)	10	54.23N/13.44E	Apr.	2003	100	4.5	9.7	Skagerrak (Ad)
Baltic Proper	Hanö Bay (HAN)	11	55.95N/15.3E	Apr.	2002	100	3.0	7.3	Natal region
	Gdansk Bay (GDA)	12	54.47N/18.57E	Mar.	2002	100	3.5	7.4	Natal region
	Gulf of Riga (RIG)	13	57.83N/22.42E	May	2002	109	5.4	5.2	Natal region
Northeast Baltic Sea	Åland Islands (ALA)	14	59.75N/22.50E	Jun.	2002	100	5.5	6.1	Natal region
	Gulf of Finland (GFI)	15	60.33N/26.17E	Jun.	2002	115	6.6	6.6	Natal region
	Swedish Bothnian Sea (SBOT)	16	61.25N/17.50E	Apr.	2002	100	2.0	5.4	Natal region
	Finnish Bothnian Sea (FBOT)	17	63.70N/21.50E	Jun.	2002	100	3.3	3.8	Natal region
	Swedish Bothnian Bay (SBB)	18	63.25N/20.50E	Jun.	2002	116	4.2	3.7	Natal region
	Finnish Bothnian Bay (FBB)	19	65.25N/23.50E	Jun.	2002	100	3.5	3.5	Natal region

[1]Average sea surface temperature/salinity in month of spawning.

[2]Specific area only given if different from natal spawning region. Juv, feeding in region as juveniles; adults feed within natal region. Ad, feeding in area as adults; juveniles feed within natal region.

standard set of control samples were analyzed for each locus on all gels, and scoring consistency between datasets was ensured by cross-referencing samples regularly during processing (also see Ruzzante et al. 2006).

ENVIRONMENTAL AND LIFE-HISTORY DATA

We were interested in investigating the potential effects of environmental and life-history factors on the overall degree of genetic isolation of each local population. Thus, environmental data were transformed prior to analysis to make them population-specific. The type of transformation varied with the environmental factor considered (see below). We considered the effect of several factors, described in terms of "environmental connectivity." One set of factors related to the physical characteristics of the sampling locations and included (1) geographical connectivity, (2) salinity, (3) temperature, (4) latitude, (5) longitude, and (6) distance to the inner Danish waters (abbreviated IDW). Another set of factors related to spawning and migratory behaviors, and included (7) spawning time, (8) number of other populations with same spawning time, (9) feeding migrations, and (10) number of other spawning groups sharing feeding ground. We describe the parameterization and rationale behind each of these factors below:

(1) *Geographic connectivity*. Pairwise geographic distances (approximated by shortest waterway distance) between sampled populations were used to calculate the mean distance between each focal population and all other population samples. The connectivity measure is inversely proportional to the degree of geographic isolation of a given local population.

(2 and 3) *Salinity and temperature*. The connectivity for salinity is best described as the mean absolute difference in salinity between the spawning ground of a sampled population during the month of spawning and that of all other sampled local populations. Note that the use of this transformation separates the correlation between salinity and longitude, allowing us to clearly distinguish between the potential effects of these factors on genetic differentiation. Temperature data were transformed using the same approach.

(4 and 5) *Latitude and longitude*. The demographic history of a species can be inferred from the structuring of neutral genetic diversity and its relationship to the spatial location of sampled populations (see Foll and Gaggiotti 2006; Kittlein and Gaggiotti 2008). Thus, we used latitude and longitude to model the potential effect of a spatial population expansion on genetic differentiation. If a population expansion took place from the West (North Sea) to the East (Baltic Sea), then we expect that local F_{ST} values will increase in this same direction because the effect of genetic drift increases as the number of founder events increases. In other words, the differentiation between the ancestral and the descendant populations increases with distance between them. This prediction was confirmed by a simulation study (see Supporting information).

(6) *Distance to IDW*. This represents the distance between each sampled population and the IDW. This factor was used to evaluate if the secondary contact zone scenario where two extant populations, one in the North Sea and the other in the Baltic Sea, progressively colonize the IDW is plausible or not. To predict the expected spatial pattern in local F_{ST} values under this scenario, we carried out a simulation study (see Supporting information).

(7 and 8) *Spawning time* and *number of populations sharing spawning time*. The time of spawning may influence the degree of genetic isolation of a local population. To model this effect, spawning time (month of peak spawning) was transformed into a "temporal connectivity" measure defined as the temporal distance between the spawning time of the focal population and the average spawning time across all populations. Month of spawning was transformed into angular data as follows. Each month is equivalent to 30° in a yearly circle and is represented as the middle angular direction, a, in the interval (15° for January, 45° for February, etc.). Based on this, the overall mean angular direction was calculated according to Zar (1996, pp. 600–601). For the samples included here, the mean angular direction was $a_{mean} = \arccos(-0.2368) \times 180/\pi = 103.7$, which corresponds to mid-April. The spawning factor for each month, i, and for the groups spawning in month i is $a_i = a_{mean} - a$ or $a_i = a - a_{mean}$, whichever is less than 180°. We also entered a factor estimating the number of populations spawning each month, because populations that reproduce at a time where few or no other populations reproduce is less likely to exchange spawners with other populations.

(9) *Feeding migrations.* Migration behavior and the distance traversed between feeding and spawning sites is likely to influence genetic structuring. If straying rates increase with migratory distance, gene flow should be proportional to migratory distance. However, because distances traversed between spawning sites and feeding areas spanning up to thousands of square kilometers may vary considerably, both within populations and between years (review in Corten 2002) we sought to parameterize this factor by tabulating whether a focal population undertakes migration beyond its geographical region or basin. Previous studies (Bekkevold et al. 2005; Jørgensen et al. 2005; Ruzzante et al. 2006) identified areas of a relatively sharp change in genetic composition, suggesting that herring populations in the area can be partitioned into the following five major groups among which gene flow seems to be restricted: (1) North Sea (samples 1–4), (2) Skagerrak (samples 5–6), (3) IDW (samples 7–10), (4) Southwestern Baltic Sea (samples 11–13), and (5) Northeastern Baltic Sea (samples 14–19). We used such partitioning to decide on the migratory behavior of the different spawning aggregations. A zero

indicated that the population did not leave its geographical region and a one indicated that it did.

(10) *Number of populations sharing feeding ground.* The proximate use of local feeding grounds may affect structure, because sharing of feeding grounds by multiple populations results in mixed-population aggregations (e.g., Ruzzante et al. 2006) and hence increases chances that individuals are socially "trapped" in nonnatal spawning migrations. This may be an especially important factor for juveniles that are naïve about spawning locations, and has been suggested to be an important parameter determining local population dynamics in herring (McQuinn 1997). The number of populations sharing feeding grounds was based on documented migration routes and feeding areas for the sampled populations (reviewed in Aro 1989 for the Baltic Sea, and in ICES 2001 for the North Sea, Skagerrak and IDW populations).

STATISTICAL METHODS

Outlier analysis

We used a modified version of the approach of Beaumont and Balding (2004) to detect loci under selection. The method is based on the idea that F_{ST} values reflect contributions from locus-specific effects, such as selection, and population-specific effects such as local effective size and immigration rates. Thus, it is possible to define F_{ST}^{ij} values that are specific to each population i and locus j and estimate them using a hierarchical Bayesian approach that models the locus and population effects using a logistic-regression model. In this modified version, implemented in the program BayeScan (Foll and Gaggiotti 2008), the posterior probability of including a locus-specific effect is estimated using a reversible jump Markov chain Monte Carlo (MCMC) approach. The outcome of this analysis was compared to the finding of Larsson et al. (2007) that suggests locus Cpa112 may be under selection. These authors used a different method, fdist (Beaumont and Nichols 1996), which uses coalescent simulations to obtain the distribution of global F_{ST} values expected under neutrality. A locus is deemed to be influenced by selection if its F_{ST} is significantly higher or lower than the expectation provided by the coalescent simulations.

Effect of environmental factors on genetic structure

We used the hierarchical Bayesian method of Foll and Gaggiotti (2006), implemented in the program GESTE, to test for the effect that different environmental factors may have on the genetic structure of herring. The method estimates individual F_{ST} values for each local population using the approach first proposed by Balding and Nichols (1995) and relates them to environmental factors using a generalized linear model. The consideration of 10 factors and their interaction leads to 2^{10} (=1024) alternative regression models and the method provides posterior probabilities for each one of them using a reversible jump MCMC approach. The model with the highest posterior probability is the one that best explains the data. We used 10 pilot runs of 1000 iterations to obtain the parameters of the proposal distributions used by the MCMC. We also used an additional burn-in of 5×10^6 iterations and a thinning interval of 50. The sample size used to obtain the estimates depended on the number of factors considered in the analyses: 60,000 with 10 factors and 30,000 with five.

The consideration of 1024 alternative models makes it difficult to interpret posterior model probabilities. The problem lies in the fact that although much more probability mass will be assigned to the most probable model, there will always be a fraction that will be allocated to models that do not explain the observed genetic differentiation pattern at all. Indeed, although the probability mass assigned to each nonprobable model is negligible (less than 0.001), the overall probability allocated to these models is nonnegligible. Thus, the posterior probability of the best model is unlikely to be very high. To overcome this problem, we adopted a strategy that consisted in first running an analysis with all 10 factors. Using its results we identified the five factors that best explain the observed genetic structure by calculating the posterior probability of including any given factor in a model. To calculate this probability for a given environmental factor, for example salinity, we would sum the posterior probability of all the models that include this factor. We then ran a second analysis that only included these five factors and that considered 32 ($=2^5$) alternative models. Based on these results, we identified the model that best explained the observed genetic structuring using the estimated posterior model probabilities. The specific effect that each factor has on genetic structuring was inferred from the estimates of the regression coefficients for the model with the highest posterior probability. This procedure was repeated for the four different analyses carried out (see below).

Results

Our objective was to fully characterize the genetic structure of herring and establish if there was a link between the outlier loci and one or more environmental factors. Thus, we adopted an approach that includes two main steps. First, we discriminated between selected and neutral loci using the outlier analysis described above. Then, we established if there were associations between environmental factors and the two kinds of loci by carrying out two types of GESTE analyses: (1) analyses that include all neutral loci plus only one outlier locus, (2) an analysis that only includes neutral loci. More precisely, the association between a particular outlier locus and one or more environmental variables was based on the comparison of the results of (1) the analysis that includes the outlier locus and all neutral loci with that of analysis (2) that only includes neutral loci. If a GESTE analysis indicates an important effect of a given environmental factor in the first analysis

(including the outlier locus) but this effect disappears when the locus is removed, then we can conclude that the environmental factor represents the selective force responsible for the outlier behavior of the locus. On the other hand, if excluding the outlier locus does not remove the effect of the environmental variable, then we can conclude that this factor has a genome-wide effect and, therefore, is not a selective factor. Note, however, that this does not mean that selection is not influencing the outlier locus because we cannot exclude the possibility that an unidentified factor that was not included in our analyses is responsible for the atypical behavior of the outlier locus. Finally, the association between environmental factors and neutral loci is obtained directly from analysis (2) that only includes neutral loci.

OUTLIER ANALYSIS

Although we used a different and considerably larger dataset and a different statistical method to detect outliers than Larsson et al. (2007), our results (see Table 2) confirmed that locus Cpa112 is probably subjected to a directional selection. Its locus-specific F_{ST} was one order of magnitude larger than that of the other loci and the posterior probability that this locus is an outlier was 1. Additionally, we identified a second locus Cpa114 that may be subject to balancing selection, based on its locus-specific F_{ST} value being lower than that of the other loci. In this latter case, the posterior probability of being an outlier was 0.998.

Having completed the first step described above, we now proceed to establish if there is an association between environmental variables and outlier as well as neutral loci.

EFFECT OF ENVIRONMENTAL FACTORS ON GENETIC STRUCTURE

As previously explained, we first carried out analyses that considered all 10 factors, and based on these results selected the five factors with the highest overall posterior probability of being included in a model. Table 3 presents these posterior probabilities for each of the 10 factors. The analysis that included locus Cpa114 and the six neutral loci assigned the highest posterior probabilities to longitude and migratory distance, followed by temperature and distance to IDW with much lower probabilities, whereas the effect of salinity ranked fifth. When Cpa112 was included the effect of salinity came out with the strongest support by far, followed by migratory distance and temperature, both with much lower scores. Finally, when only neutral loci were included, migratory distance had the highest posterior probability, followed by longitude and distance to IDW, while the effects of salinity and temperature decreased considerably. We considered that the best five factors were those included among the five best in at least two of the three analyses. These were longitude, salinity, feeding migration, temperature, and distance to the IDW.

Table 4 presents the results of analyses that considered the five best factors. We present results for the five models with the highest posterior probability because the remaining 27 models had negligible posterior probabilities. Posterior model probabilities changed substantially depending on whether one of the outlier loci (Cpa114 and Cpa112) was included in the dataset (Table 4). When locus Cpa114 was included, the best model included migratory distance only, whereas the second best included salinity and longitude. The analysis that included locus Cpa112 identified the model that included salinity only as the best model, with a posterior probability much larger than that of the second-best model, which included temperature and migratory distance. Finally, the analysis that included only neutral loci, assigned the highest posterior probability to the model that included migratory

Table 2. Results of the outlier analysis showing that the two loci Cpa112 and Cpa114 are highly likely to be outliers and thereby not selectively neutral. α is the locus specific effect and Pr(α≠0) is the posterior probability that the model does not include the locus-specific effect (i.e., the probability that the locus is not an outlier).

Locus	α	Pr(α≠0)	F_{ST}
Cha1017	−0.476	0.517	0.00318
Cha1020	0.389	0.558	0.00628
Cha1027	−0.270	0.307	0.00402
Cha1202	−0.346	0.379	0.00365
Cpa101	−0.320	0.370	0.00379
Cpa111	0.424	0.471	0.00697
Cpa112	2.570	1.000	0.05560
Cpa114	−0.981	0.998	0.00181

Table 3. Sum of posterior probabilities of models that include a given factor. GESTE analyses included all 10 factors. Bold value indicates factor with highest score.

Factor	Sum of posterior probabilities		
	Cpa114 and neutral	Cpa112 and neutral	neutral only
Longitude	**0.463**	0.166	0.335
Latitude	0.116	0.084	0.119
Salinity	0.195	**0.552**	0.184
Feeding migrations	0.423	0.266	**0.412**
Temperature	0.252	0.252	0.148
Geographical connectivity	0.193	0.137	0.174
No. of populations spawning same time	0.141	0.095	0.122
Spawning time	0.089	0.056	0.109
No. of spawning groups sharing feeding ground	0.100	0.069	0.100
Distance to IDW	0.235	0.184	0.216

Table 4. Posterior probabilities of the five most probable models for the analyses including the five factors that obtained an overall posterior probability of at least 0.25 in either one of the two analyses that considered all 10 factors. (A) locus Cpa114 and neutral loci, (B) locus Cpa112 and neutral loci, (C) neutral loci only. Models are listed in decreasing order of posterior probabilities.

Model	Pr	Factors included
(A)		
5	0.256	Migratory distance
4	0.163	Salinity, longitude
10	0.108	Temperature, longitude
2	0.077	Longitude
1	0.062	Null
(B)		
3	0.353	Salinity
13	0.140	Temperature, migratory distance
4	0.102	Salinity, longitude
1	0.074	Null
7	0.054	Migratory distance, salinity
(C)		
5	0.286	Migratory distance
1	0.140	Null
4	0.125	Salinity, longitude
2	0.092	Longitude
10	0.058	Temperature, longitude

Table 5. Posterior estimates of regression parameters for the model with the highest posterior probability when five factors are considered. (A) locus Cpa114 and neutral loci, (B) locus Cpa112 an neutral loci, (C) neutral loci only.

Regression coefficient	Factor	Mean	Mode	95% HPDI
(A)				
α_0	Constant	−5.780	−5.740	[−6.18; −5.40]
α_3	Migrations	0.566	0.547	[0.194; 0.947]
σ^2	–	0.505	0.366	[0.160; 1.020]
(B)				
α_0	Constant	−4.710	−4.73	[−5.00; −4.430]
α_2	Salinity	−0.429	−0.437	[−0.696; −0.158]
σ^2	–	0.286	0.219	[0.113; 0.537]
(C)				
α_0	Constant	−5.63	−5.62	[−6.020; −5.240]
α_3	Migrations	0.509	0.486	[0.150; 0.888]
σ^2	–	0.485	0.365	[0.145; 0.935]

distance only, whereas the second best was the null model, which excludes all factors.

It is worth noting that the results of the analyses that considered all 10 factors (Table S1) were very similar to those presented in Table 4. More precisely, the best model remained the same, and there were few changes in the ranking of the remaining four best models. Additionally, the estimates of the regression parameters for the models with the highest posterior probability (Table S2) were almost identical to those presented in Table 5.

Several general remarks can be made based on these results. The comparison of the results of the analysis that include locus Cpa112 with those of the analysis that considers only neutral markers clearly indicates an association between this locus and salinity. More precisely, the posterior probability of the model that includes salinity only is the highest when locus Cpa112 is included but drops to third place after the null model when this outlier locus is excluded. This observation, taken together with the stronger than expected genetic structuring at this locus, gives support to the hypothesis that it is influenced by directional selection. On the other hand, the comparison of the results of the analysis that includes locus Cpa114 with those of the analysis that considers only neutral markers suggests that none of the environmental factors considered represents a selective force responsible for the outlier behavior of this locus. Although this does not mean that this locus is not influenced by selection, we note that the weaker than expected genetic structuring together with its lack of association with environmental factors suggest that an atypical mutation rate, very different from that of typical neutral loci, could be a plausible alternative explanation. In terms of the structuring of neutral genetic variation, the analysis that includes only neutral markers indicates that feeding migrations have an important effect. Although minimizing the effect of selection (and/or mutation) by excluding both outlier loci further enhances our ability to discriminate the effect of migratory distance, its importance in structuring genetic diversity was already apparent in the two analyses that included one of the outlier loci because models including this factor ranked second (when Cpa112 was included) or first (when Cpa114 was included). Finally, the demographic history of the species seems to have only a minor influence, because models including longitude or distance to IDW had low posterior probabilities.

The specific effect of each factor on genetic structuring can be inferred from the estimates of the regression coefficients under the best model. As mentioned before, the results were almost identical, whether we consider all 10 factors or only the five best (compare Table 5 of main text and Table S2). When outlier locus Cpa112 was included in the analyses, the genetic isolation of local populations decreased (i.e., local F_{ST} decreased) as the difference in salinity between the local population and all other populations increased (negative α_2 in Table 5C). To interpret this result, it is necessary to note that, as explained before, the effect of salinity was introduced by calculating a connectivity measure based on the absolute difference in salinity between each focal population and the average salinity across the whole range. Thus, the higher salinity-connectivity values correspond to areas that have either high or low salinity, which are, respectively, the North Sea and

the Baltic Sea. On the other hand areas with intermediate salinity will have low salinity-connectivity values. The negative regression coefficient means that local population F_{ST} decreases as salinity-connectivity values increase. In other words, populations that inhabit areas with high (North Sea) or low salinity (Baltic Sea) are well connected genetically, whereas those that inhabit areas with intermediate salinities (IDW) are genetically isolated. This suggests that gene flow from populations inhabiting either high or low salinity conditions into populations inhabiting intermediate salinity areas is rare and that most exchanges occur among populations inhabiting similar salinity conditions. It is worth noting that the amount of variation that remains unexplained by the regression model, which is measured by the estimates of σ^2 (see Foll and Gaggiotti 2006), is lowest when locus Cpa112 is included in the analysis. This provides further evidence that the effect of selection is important. In terms of the factors that influence neutral genetic diversity, only the model with feeding migrations had a higher posterior probability than the null model. The positive regression coefficient indicates that the genetic isolation of populations increased with migratory distance (positive α_3 in Table 5C), a result that suggests that long-range migration does not necessarily lead to gene flow.

Discussion

Our study illustrates the complex interplay between environmental factors, behavioral/physiological traits and demographic history that are likely to influence the structuring of genetic diversity in pelagic fish species. Our approach allowed us to disentangle the individual potential effects of each of these factors on the genetic structure and can in principle be applied to any abundant and widely distributed aquatic or terrestrial species. Importantly, such an approach can be used to generate a specific hypothesis-testing framework that can be examined empirically. As a first necessary step we discriminated between neutral and selected variation. Such distinction allowed us to establish that the two most important factors were (1) an environmental factor, salinity on spawning sites and (2) a behavioral trait, migration with philopatry. The genetic signal left by the demographic history of herring, on the other hand, seems to have been largely eroded, which is not surprising given that the large reproductive potential and presumed enormous local effective population sizes of pelagic fish act to minimize the effect of genetic drift (cf. Nei et al. 1975). Note that although this species is subject to fishing, theoretical work indicates that harvesting is likely to play only a minor role in shaping genetic structure of species with very large N_e and several overlapping generations (Gaggiotti and Vetter 1999). Although we identified two outlier loci, only one of them, locus Cpa112, was associated to an environmental factor (salinity) and exhibited a spatial pattern of genetic structuring that is consistent with directional selection. The second outlier locus, Cpa114, was not associated with any environmental factor, and its spatial genetic structure was more homogeneous than expected under neutrality. Although these two observations do not rule out the effect of (balancing) selection, they do suggest that we cannot exclude a very atypical mutation rate as another potential explanation for the observed genetic structure at this locus.

The approach we used to establish an association between outlier loci and environmental factors includes a first step based on an outlier analysis built on the idea that F_{ST} is influenced by genome-wide effects and locus-specific effects, the latter being the result of selective pressures. The ability to distinguish between these two effects improves as the number of markers increases. In this regard, we have to note that we used only eight microsatellite loci, and two of them were identified as outliers. Ideally, the total number of loci should be larger but we note that the outlier behavior of one of the loci, Cpa112, was already observed by a previous analysis (Larsson et al. 2007) using a different dataset. Additionally, the removal of the outlier loci radically changed the results of GESTE analyses, whereas removing neutral loci did not lead to important changes (see Table S3).

SALINITY

A previous study (Bekkevold et al. 2005) showed that patterns of reproductive isolation in herring covaried with salinity differences among spawning locations, independently of geographical distance. Our study confirms that salinity has an important effect on the genetic structuring of this species, but suggests that its effect is locus-specific and may represent a selective pressure. More precisely, outlier locus Cpa112 seems to be associated with a genomic region involved in (low) salinity adaptation. When this locus was included in the GESTE analysis, the model including only salinity had the highest posterior probability (see Table 4), but when it was removed, this posterior probability decreased pronouncedly being lower than that of the null model (which only has the constant term). It is worth noting that the outlier behavior of locus Cpa112 was identified by the study of Larsson et al. (2007), who suggested that it could be influenced by directional selection. This would explain why the F_{ST} value at this locus was much higher than that of the remaining seven microsatellite loci. Note that the effect of including locus Cpa112 was consistent across populations, increasing local F_{ST} across the board (see the last two columns in Table S4). The genomic position of Cpa112 remains to be determined, and the observed signal may be attributed either to linkage with a gene under divergent selection (hitchhiking effect) or selection on the locus itself. The hypothesis that Cpa112 is associated with salinity adaptation is further supported by results for a population from a brackish Skagerrak fjord, which for Cpa112 exhibits allele frequencies similar to populations in IDW, but close genetic relationships with other Skagerrak populations for other

microsatellite loci (C. Andre, unpubl. data). Moreover, the results suggest that gene flow from populations inhabiting either high or low salinity conditions into populations inhabiting intermediate salinity areas is rare. Selection due to salinity is probably high during early stages of the herring's life cycle. Eggs are fairly tolerant to changes in salinity of the surrounding water, but the lower tolerable limit changes with increasing distance to the marine transition zone toward the North Sea, whereas Northeastern Baltic herring eggs tolerate almost fresh water (e.g., Klinkhardt 1996). Our results hence suggest that the contact zone does not merely constitute a hybrid-sink (cf. Barton 1980). Rather, contact zone populations are likely physiologically adapted to their local environment, as also implied for cod in the same area, based on population-specific fertilization and egg buoyancy capacities at different salinities (Nissling and Westin 1997).

PHILOPATRY

Our expectation was that long migration distances would lead migrating herring past nonnatal spawning areas and increase the likelihood of straying. Contrary to this, migration distance showed positive regression coefficients, suggesting that populations exhibiting migration beyond their natal region, also tended to be more genetically distinct than less migratory populations. Thus, long-range migration does not lead to gene flow per se, and the extent of natal homing hence seems to play an important role in structuring genetic diversity across populations. Moreover, the evolution of homing invokes an advantage to reproducing at the natal site, due to either a local adaptation to that particular site or to the ascertaining of a predictable spawning habitat (Travis and Dytham 1999). A scenario could be that migration leads to increased risk of straying into environments where the phenotype is maladapted (Hendry 2004; Bolnick and Nosil 2007). This could drive selection for natal homing to be stronger in populations that undertake long-range migrations outside their geographical regions or basins, than in populations migrating shorter distances. However, such mechanisms are yet unknown for herring, as well as for most marine fish.

Isolation by distance had no significant structuring effect per se. Although both isolation by distance and natal homing are commonly assumed to contribute to the maintenance of population structure in marine and anadromous fish with large migratory potential (e.g., Palstra et al. 2007), attempts to assess their relative importance remain scarce (see Dionne et al. 2008 for a recent study). Our approach is novel as it allowed separation of the effects of homing and isolation by distance by combining analytically detailed population genetic information with the large body of data on herring migrations obtained using both phenotypic (see ICES 2001) and genetic marker-based (Ruzzante et al. 2006) approaches.

REPRODUCTIVE TIMING

In some organisms reproductive timing has a heritable component and acts as a barrier to gene flow among divergent populations (e.g., Miyatake 2002; Hendry and Day 2005; Antonovics 2006; O'Malley et al. 2008). Our analysis indicated that this was not the case in herring, as none of the factors related to spawning time showed measurable effects on structure in concordance with reports that spawning time can switch within local population components (e.g., Bekkevold et al. 2007). Spawning time thus appears to exhibit strong phenotypic plasticity in herring, in contrast to some salmonid species, where it is found to have a strong heritable component (Fillatre et al. 2003; O'Malley et al. 2008). In herring, such plasticity seemingly enables an opportunistic response to environmental variability, as is for example, also implicated in many avian species (Pulido 2007).

DEMOGRAPHIC HISTORY

We aimed to distinguish between two alternative scenarios for the demographic history of herring: either a progressive colonization from the North Sea in an eastward direction, or a secondary contact zone with two ancestral populations in, respectively, the North Sea and the Baltic Sea that progressively colonize the transition zone.

Previous studies did not discriminate between neutral and selected genetic variation as we have done here, and therefore could not evaluate to what extent the observed clinal variation may represent a primary contact zone shaped by divergent selection, or a secondary contact zone reflecting the demographic history of herring. Having first established that selection is a likely explanation for the genetic structuring of variation at locus Cpa112, we addressed the issue of demographic history by focusing on neutral markers only (i.e., those that exclude both Cpa112 and Cpa114). The factors used to take into account the effect of demographic history (longitude and distance to the IDW) had limited effects on genetic structuring compared to salinity and migration distance, but they nevertheless ranked second and third in the analyses that only included neutral markers, and fourth and fifth in the other analyses (cf. Table 3). In case of a progressive colonization from the North Sea, we expect a steady increase in F_{ST} as we move from the North Sea toward the Baltic Sea because of the repeated founder events that take place as the eastward colonization of the habitat proceeds (Fig. S1A). The observed pattern is different to such predictions, with high F_{ST}s for intermediate longitudes, and a progressive decrease toward both the North Sea and Baltic Sea; a pattern consistent with the secondary contact zone scenario (Fig. S1B). The indication of a secondary contact zone may reflect that Baltic populations are the result of one major founding event, as hypothesized for a range of marine organisms in the Baltic (review in Johannesson and Andre 2006). North Sea herring populations were likely founded by range expansion from Atlantic populations with the creation of the English Channel after 7500BP

(Lambeck 1995). The biogeographical history of Baltic populations is more uncertain. The Baltic Sea started turning saline after ~8900BP when oceanic waters entered from the North Sea via the Skagerrak and IDW, with a northeasterly progression of saline conditions over the next ~2000 years (Sohlenius et al. 2001). However, during this period, salinity levels varied greatly over both time and space in response to repeated glaciation and melting dynamics. The Gdansk Deep, for example, remained saline for long periods during which freshwater conditions prevailed throughout shallower parts of the Baltic. It is, therefore, conceivable that herring persisted and underwent adaptive divergence in one or more large brackish refugia completely isolated from populations in the saline North Sea. Subsequent development of stable brackish conditions throughout the Baltic at ~6500 years BP may then have led to a swift range expansion throughout the Baltic and secondary contact with North Sea populations.

In combination with the inference about salinity-related divergent selection and our result that gene flow is restricted between the transition zone and adjacent populations, it can be hypothesized that hybrid inferiority and potentially reinforcement through homing behavior act to maintain the genetic cline through a secondary contact zone.

Conclusions

In populations with large effective population sizes, the difference in genetic divergence for neutral versus selected loci is expected to be particularly high due to a relatively large effect of selection in comparison with genetic drift (e.g., Allendorf et al. 2008). The inclusion of loci under (hitchhiking) selection is thus likely to have a large effect on analyses of structure and demography. More specifically, loci under divergent selection can inflate estimates of genetic isolation (e.g., Nielsen et al. 2006), whereas loci under stabilizing selection can decrease it. Nonetheless, including and comparing information for neutral and selected loci (including candidate gene markers) can yield important information when attempting to identify the processes that potentially affect the spatial distribution of genetic resources in an organism (Hemmer-Hansen et al. 2007; Larsen et al. 2008), and for resolving conservation and management issues (see Hauser and Carvalho 2008 for a review). The present study shows that such an approach provides a powerful tool for disentangling the effects of evolutionary, demographic, and environmental factors influencing the genetic structure of natural populations. The approach can moreover be useful for designing targeted population genomic studies (also see Hemmer-Hansen et al. 2007). Our study also contributes empirical evidence that microsatellite-based estimates of weak structure should not be interpreted to indicate that gene flow is high, or that adaptive divergence is impeded in organisms with large effective population sizes, that may or may not have attained migration–drift equilibrium (Conover et al. 2006; Larsen et al. 2007; Mäkinen et al. 2008).

ACKNOWLEDGMENTS

The comments made by two anonymous reviewers greatly improved the original manuscript. The collaboration that led to this article was partly funded by grant 1034 from the programme CONGEN, European Science Foundation. The work was partly based on research carried out in the HERGEN project (http://www.hull.ac.uk/hergen/), funded by the European Union within the 5th framework program, and the efforts of the joint consortium are acknowledged.

LITERATURE CITED

Allendorf, F. W., P. R. England, G. Luikart, P. A. Ritchie, and N. Ryman. 2008. Genetic effects of harvest on wild animal populations. Trends Ecol. Evol. 23:327–337.

Antonovics, J. 2006. Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. Heredity 97:33–37.

Balding, D. J., and R. A. Nichols. 1995. A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. Genetica 96:3–12.

Barton, N. H. 1980. The hybrid sink effect. Heredity 44:277–278.

Beaumont, M. A., and D. J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. 13:969–980.

Beaumont, M. A., and R. A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond. B 263:1619–1626.

Bekkevold, D., C. Andre, T. G. Dahlgren, L. A. W. Clausen, E. Torstensen, H. Mosegaard, G. R. Carvalho, T. B. Christensen, E. Norlinder, and D. E. Ruzzante. 2005. Environmental correlates of population differentiation in Atlantic herring. Evolution 59:2656–2668.

Bekkevold, D., L. A. W. Clausen, S. Mariani, C. Andre, T. B. Christensen, and H. Mosegaard. 2007. Divergent origins of sympatric herring population components determined using genetic mixture analysis. Mar. Ecol. Prog. Ser. 337:187–196.

Bolnick, D. I., and P. Nosil. 2007. Natural selection in populations subject to a migration load. Evolution 61:2229–2243.

Conover, D. O., L. M. Clarke, S. B. Munch, and G. N. Wagner. 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. J. Fish Biol. 69:21–47.

Corten, A. 2002. The role of "conservatism" in herring migrations. Rev. Fish Biol. Fish. 11:339–361.

Dionne, M., F. Caron, J. J. Dodson, and L. Bernatchez. 2008. Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Mol. Ecol. 10:2382–2396.

Fillatre, E. K., P. Etherton, and D. D. Heath. 2003. Bimodal run distribution in a northern population of sockeye salmon (*Oncorhynchus nerka*): life history and genetic analysis on a temporal scale. Mol. Ecol. 12:1793–1805.

Foll, M., and O. E. Gaggiotti. 2006. Identifying the environmental factors that determine the genetic structure of populations. Genetics 174:875–891.

———. 2008. Estimating selection with different markers and varying demographic scenarios: a Bayesian perspective. Genetics 180:977–993.

Gaggiotti, O. E., and R. D. Vetter. 1999. Effect of life history strategy, environmental variability, and overexploitation on the genetic diversity of pelagic fish populations. Can. J. Fish. Aquat. Sci. 56:1376–1388.

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

HELCOM. 1996. Third periodic assessment of the state of the marine environment of the Baltic Sea, 1989–1993, background document. Proc. Baltic Sea Environ. Symp. 64B:l–252.

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

Hendry, A. P. 2004. Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. Evol. Ecol. Res. 6:1219–1236.

Hendry, A. P., and T. Day. 2005. Population structure attributable to reproductive time: isolation by time and adaptation by time. Mol. Ecol. 14:901–916.

Iles, T. D., and M. Sinclair. 1982. Atlantic herring—stock discreteness and abundance. Science 215:627–633.

Johannesson, K., and C. Andre. 2006. Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. Mol. Ecol. 15:2013–2029.

Jorgensen, H. B. H., M. M. Hansen, D. Bekkevold, D. E. Ruzzante, and V. Loeschcke. 2005. Marine landscapes and population genetic structure of herring (*Clupea harengus* L.) in the Baltic Sea. Mol. Ecol. 14:3219–3234.

Kittlein, M., and O. E. Gaggiotti. 2008. Interactions between environmental factors can preclude the detection of isolation by distance patterns: a case study of *Ctenomys rionegrensis* in Uruguay. Proc. R. Soc. Lond. B 275:2633–2638.

Klinkhardt, M. 1996. Der Hering. Spektrum Akademie Verlag, Heidelberg.

Lambeck, K. 1995. Glacial isostasy and water depths in the late Devensian and Holocene on the Scottish shelf west of the Outer-Hebrides. J. Quat. Sci. 10:83–86.

Larsen, P. F., E. E. Nielsen, T. D. Williams, J. Hemmer-Hansen, J. K. Chipman, M. Kruhøffer, P. Grønkjær, S. G. George, L. Dyrskjøt, and V. Loeschcke. 2007. Adaptive differences in gene expression in European flounder (*Platichthys flesus*). Mol. Ecol. 16:4674–4683.

Larsen, P. F., E. E. Nielsen, T. D. Williams, and V. Loeschcke. 2008. Intraspecific variation in expression of candidate genes for osmoregulation, heme biosynthesis and stress resistance suggests local adaptation in European flounder (*Platichthys flesus*). Heredity 101:247–259.

Larsson, L. C., L. Laikre, S. Palm, C. Andre, G. R. Carvalho, and N. Ryman. 2007. Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. Mol. Ecol. 16:1135–1147.

Mäkinen, H. S., M. Cano, and J. Merila. 2008. Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Mol. Ecol. 17:3565–3582.

Mariani, S., W. F. Hutchinson, E. M. C. Hatfield, D. E. Ruzzante, E. J. Simmonds, T. G. Dahlgren, C. Andre, J. Brigham, E. Torstensen, and G. R. Carvalho. 2005. North Sea herring population structure revealed by microsatellite analysis. Mar. Ecol. Prog. Ser. 303:245–257.

McPherson, A. A., P. T. O'Reilly, T. L. McParland, M. W. Jones, and P. Bentzen. 2001. Isolation of nine novel tetranucleotide microsatellites in Atlantic herring (*Clupea harengus*). Mol. Ecol. Notes 1:31–32.

McPherson, A. A., P. T. O'Rielly, and C. T. Taggart. 2004. Genetic differentiation, temporal stability, and the absence of isolation by distance among Atlantic herring populations. Trans. Am. Fish. Soc. 133:434–446.

McQuinn, I. H. 1997. Metapopulations and the Atlantic herring. Rev. Fish Biol. Fish. 7:297–329.

Miyatake, T. 2002. Pleiotropic effect, clock genes, and reproductive isolation. Popul. Ecol. 44:201–207.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

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

Nielsen, E. E., P. H. Nielsen, D. Meldrup, and M. M. Hansen. 2004. Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea. Mol. Ecol. 13:585–595.

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

Nissling, A., and L. Westin. 1997. Salinity requirements for successful spawning of Baltic and Belt Sea cod and the potential for cod stock interactions in the Baltic Sea. Mar. Ecol. Prog. Ser. 152:261–271.

Olsen, J. B., C. J. Lewis, E. J. Kretschmer, S. L. Wilson, and J. E. Seeb. 2002. Characterization of 14 tetranucleotide microsatellite loci derived from Pacific herring. Mol. Ecol. Notes 2:101–103.

O'Malley, K. G., and M. A. Banks. 2008. A latitudinal cline in the Chinook salmon (*Onchorhynchus tsawytscha*) Clock gene: evidence for selection on PolyQ length variants. Proc. R. Soc. Lond. B 275:2813–2821.

Palstra, F. P., and D. E. Ruzzante. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Mol. Ecol. 17:3428–3447.

Palstra, F. P., M. F. O'Connell, and D. E. Ruzzante. 2007. Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. Mol. Ecol. 16:4504–4522.

Palumbi, S. R. 1994. Genetic-Divergence, reproductive isolation, and marine speciation. Annu. Rev. Ecol. Syst. 25:547–572.

Pulido, F. 2007. Phenotypic changes in spring arrival: evolution, phenotypic plasticity, effects of weather and condition. Clim. Res. 35:5–23.

Ruzzante, D. E., S. Mariani, D. Bekkevold, C. Andre, H. Mosegaard, L. A. W. Clausen, T. G. Dahlgren, W. F. Hutchinson, E. M. C. Hatfield, and E. Torstensen, et al. 2006. Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. Proc. R. Soc. Lond. B 273:1459–1464.

Sohlenius, G., K. C. Emeis, E. Andren, T. Adren, and A. Kohly. 2001. Development of anoxia during the Holocene fresh-brackish water transition in the Baltic Sea. Marine Geol. 177:221–242.

Travis, J. M. J., and C. Dytham. 1999. Habitat persistence, habitat availability and the evolution of dispersal. Proc. R. Soc. Lond. B 266:723–728.

Turner, T. F., M. J. Osborne, G. R. Moyer, M. A. Benavides, and D. Alo. 2006. Life history and environmental variation interact to determine effective population to census size ratio. Proc. R. Soc. Lond. B 273:3065–3073.

Zar, J. H. 1996. Biostatistical analysis. 3rd edn. Prentice-Hall. Englewood Cliffs, New Jersey.

Associate Editor: S. Gandon

Supporting Information

The following supporting information is available for this article:

Figure S1. Comparison of SPLATCHE simulation results with the observed longitudinal pattern of local F_{ST} variation.

Table S1. Posterior probabilities of the 10 most probable models for the analyses including all 10 factors.

Table S2. Posterior estimates of regression parameters for the model with the highest posterior probability when all 10 factors are considered.

Table S3. Posterior estimates of regression parameters for the model with the highest posterior probability when five factors are considered.

Table S4. Mode and 95% HPDIs (in parentheses) of local population F_{ST} estimates for all four datasets.

Supporting Information may be found in the online version of this article.
(This link will take you to the article abstract).