Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean

Sangeet Lamichhaney[a,1], Angela P. Fuentes-Pardo[b,1], Nima Rafati[a], Nils Ryman[c], Gregory R. McCracken[b], Christina Bourne[d], Rabindra Singh[e], Daniel E. Ruzzante[b], and Leif Andersson[a,f,g,2]

[a]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, 752 36 Uppsala, Sweden; [b]Department of Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada; [c]Department of Zoology, Stockholm University, 106 91 Stockholm, Sweden; [d]Fisheries and Oceans Canada, Northwest Atlantic Fisheries Centre, St John's, Newfoundland A1C 5X1, Canada; [e]Fisheries and Oceans Canada, St. Andrews Biological Station, St. Andrews, New Brunswick E5B 2L9, Canada; [f]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden; and [g]Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX 77843

Contributed by Leif Andersson, February 24, 2017 (sent for review October 27, 2016; reviewed by Dorte Bekkevold and Craig R. Primmer)

Atlantic herring is an excellent species for studying the genetic basis of adaptation in geographically distant populations because of its characteristically large population sizes and low genetic drift. In this study we compared whole-genome resequencing data of Atlantic herring populations from both sides of the Atlantic Ocean. An important finding was the very low degree of genetic differentiation among geographically distant populations (fixation index = 0.026), suggesting lack of reproductive isolation across the ocean. This feature of the Atlantic herring facilitates the detection of genetic factors affecting adaptation because of the sharp contrast between loci showing genetic differentiation resulting from natural selection and the low background noise resulting from genetic drift. We show that genetic factors associated with timing of reproduction are shared between genetically distinct and geographically distant populations. The genes for thyroid-stimulating hormone receptor (*TSHR*), the SOX11 transcription factor (*SOX11*), calmodulin (*CALM*), and estrogen receptor 2 (*ESR2A*), all with a significant role in reproductive biology, were among the loci that showed the most consistent association with spawning time throughout the species range. In fact, the same two SNPs located at the 5′ end of *TSHR* showed the most significant association with spawning time in both the east and west Atlantic. We also identified unexpected haplotype sharing between spring-spawning oceanic herring and autumn-spawning populations across the Atlantic Ocean and the Baltic Sea. The genomic regions showing this pattern are unlikely to control spawning time but may be involved in adaptation to ecological factor(s) shared among these populations.

genetic adaptation | Atlantic herring | parallel evolution | reproductive strategies | whole-genome resequencing

Widely dispersed and abundant species generally exhibit populations inhabiting divergent habitats. Such populations need to adapt to local biotic and abiotic factors, a process that results in higher fitness in the local environment and leads to genetic and phenotypic differentiation among subpopulations (1, 2). In addition to such local adaptation, if a trait responds to similar forces of natural selection independently across multiple populations or species, parallel evolution will lead to similar adaptive changes among geographically distant populations (3). Such parallel adaptation may be caused by convergent evolution or by the sharing of similar (or identical) genetic changes across populations (4–6). Identification of the genetic basis for ecological adaptation is a fundamental goal in evolutionary biology (7), and current technological advances in population-scale high-throughput sequencing provide powerful tools to explore these processes at a genomic scale (8). However, many adaptive traits are expected to have a highly polygenic background (9) in which genes of small effect are hard to detect with traditional genome scans, and adaptive changes can often be confounded with demographic history effects such as population structure and genetic drift (10), making the dissection of such genetic differentiations challenging.

Our recent population-scale genomic study on Atlantic herring has demonstrated that this species is an ideal model for the detection of signatures of selection (11, 12). The large effective population size (Atlantic herring is one of the most abundant fish species on earth) probably combined with gene flow between populations results in extremely low levels of genetic differentiation at selectively neutral loci across populations exposed to different ecological conditions. These factors allowed us to identify about 500 independent loci associated with local adaptation as regards the colonization of the brackish Baltic Sea and timing of reproduction in Northeast (NE) Atlantic herring populations (11).

Atlantic herring is a schooling pelagic fish distributed throughout the North Atlantic Ocean and adjacent waters, including the Baltic Sea (Fig. 1). The population structure of Atlantic herring is considered to be one of the most complex of any marine fish, and there is a long history of attempts to describe it (13). Traditionally, herring

Significance

Identification of genetic changes that allow a species to adapt to different environmental conditions is an important topic in evolutionary biology. In this study we analyzed whole-genome resequencing data of Atlantic herring populations from both sides of the Atlantic Ocean and identified a number of loci that show consistent associations with spawning time (spring or autumn). Several of these loci, such as thyroid-stimulating hormone receptor (*TSHR*), have a well-established role in reproductive biology, whereas others have never been implicated in controlling reproduction. Genetic variants associated with adaptation to spring or autumn spawning are shared to a large extent among populations across the Atlantic Ocean and the Baltic Sea, providing evidence for parallel adaptive evolution.

Author contributions: D.E.R. and L.A. conceived the study; S.L., A.P.F.-P., N. Rafati, and N. Ryman performed research; G.R.M., C.B., and R.S. contributed critical samples for the study; and S.L., A.P.F.-P., D.E.R., and L.A. wrote the paper.

Reviewers: D.B., Technical University of Denmark; and C.R.P., University of Turku.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The Illumina reads reported in this article have been submitted to the Short Reads Archive (https://www.ncbi.nlm.nih.gov/sra) (accession no. PRJNA338612).

[1]S.L. and A.P.F.-P. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: leif.andersson@imbim.uu.se.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1617728114/-/DCSupplemental.

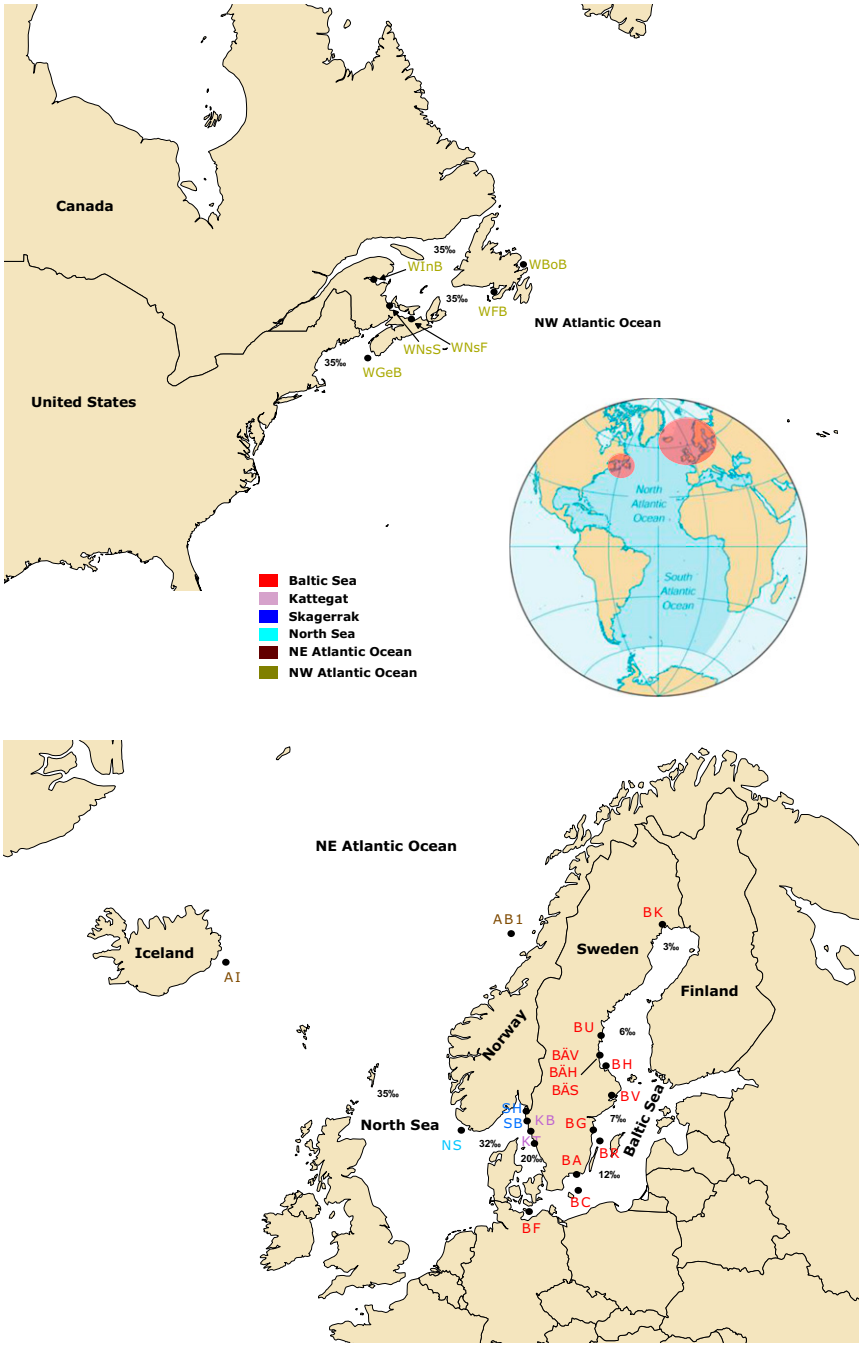

Fig. 1. Sampling sites. Geographic location of all population samples. Abbreviations for localities are given in Table 1. The relative locations of populations sampled from the NW and NE Atlantic Ocean are indicated in the inserted globe. Maps were created using ArcGIS (Esri).

stocks have been described based on morphology and life history traits such as spawning time and location (13–15). Populations are known to spawn in different seasons, with some spawning in the spring, others in the autumn, and still others at intermediate times. In each case, spawning generally takes place over a protracted period of a few weeks. The optimal spawning time is generally linked to environmental conditions associated with plankton blooms (16). Our recent study revealed highly significant genetic differentiation between spring- and autumn-spawning herring in the NE Atlantic, with some of the loci involved in this differentiation likely controlling the timing of reproduction (11).

Our previous genome-scale studies of Atlantic herring were restricted to population samples from the NE Atlantic, but this species is ecologically important throughout the North Atlantic. In fact, herring support a commercially important fishery in the Northwest (NW) Atlantic (17), where the species is recognized for the complexity and plasticity of its stocks (13, 18). Similar to the NE Atlantic populations, herring in the NW Atlantic undergo north–south and inshore–offshore migrations for feeding and reproduction (19), with spawning taking place mostly during spring and autumn (20) from Cape Cod to northern Newfoundland (14). Previous genetic studies based on a small number of microsatellite markers reported weak but significant genetic structuring between NW and NE Atlantic populations as well as among spawning aggregations within the NW Atlantic (21, 22).

The presence of spring- and autumn-spawning herring populations on both sides of the North Atlantic provides an exceptional opportunity to explore whether the same or similar genetic factors associated with spawning time are shared by geographically

distant populations. In this study, we present the results of whole-genome resequencing of six NW Atlantic populations (three spring and three autumn spawners) and compare their genetic architecture with that of NE Atlantic populations. We demonstrate that genetic factors associated with the timing of reproduction are shared to a large extent between herring populations from the NW and NE Atlantic.

Results

Whole-Genome Resequencing. Whole-genome resequencing of pooled DNA of 40–50 individuals per location was conducted for six population samples from the NW Atlantic: two populations from Newfoundland, three from the Gulf of St Lawrence, and one from the Scotian Shelf (Fig. 1 and Table 1). Each library had an ∼30–50× depth of coverage. These data were compared with pooled sequence data from 19 populations of Baltic Sea/NE Atlantic herring (Fig. 1) and one population of Pacific herring from our previous study (11); the Pacific herring data were used as an outgroup in the phylogenetic analysis and allowed us to determine the ancestral state for candidate causal mutations. In addition, 27 individual samples of spring- and autumn-spawning herring from both sides of North Atlantic and six individual samples of Pacific herring were sequenced to ∼10× depth of coverage to characterize haplotypes showing genetic differentiation. These sequences were aligned to the reference herring genome (11), and SNP calling was conducted using an in-house rigorous quality-filtering pipeline (*Materials and Methods*) to identify 8.9 million SNPs that were polymorphic in at least one population (including Pacific herring).

Phylogeny and Population Genetics. In agreement with the results from our previous study (11), the neighbor-joining tree generated using 1.2 million high-quality SNPs (*Materials and Methods*) revealed a large genetic distance between the Pacific and all 25 Atlantic herring populations (Fig. 2, *Left*). The Atlantic herring populations in general clustered according to their geographic origin (Fig. 2, *Right*). The populations formed three major groupings: (*i*) Atlantic herring from NW Atlantic and NE Atlantic (populations AB1 and AI); (*ii*) Atlantic herring from the North Sea, Skagerrak, and Kattegat; and (*iii*) spring-spawning herring from the Baltic Sea. Two populations of autumn-spawning herring from the Baltic Sea (BÄH and BF) deviated from this pattern and did not cluster with spring-spawning herring from the Baltic Sea. Two populations (NS and WFB) stood out with relatively long branch lengths; a careful examination of the data from these populations did not indicate that these long branch lengths were caused by technical issues. Among the six NW Atlantic populations, spring and autumn spawners formed their own clusters, indicating that populations spawning in different seasons are genetically distinguishable. NW Atlantic autumn-spawning herring were more closely related to spring-spawning herring from the NW Atlantic than to autumn-spawning herring from the NE Atlantic.

The average fixation index (F_{ST}) among the 25 populations of Atlantic and Baltic herring was as low as 0.026, a value that drops

Table 1. Population samples of herring used in the study

Locality	Sample	*n* (*n2*)	Position		Salinity, ‰	Date	Spawning season‡
Baltic Sea							
Gulf of Bothnia (Kalix)*	BK	47	N65°52′	E22°43′	3	1980 June 29	Summer
Bothnian Sea (Hudiksvall)*	BU	100	N61°45′	E17°30′	6	2012 April 19	Spring
Bothnian Sea (Gävle)*	BÄV	100	N60°43′	E17°18′	6	2012 May 7	Spring
Bothnian Sea (Gävle)*	BÄS	100	N60°43′	E17°18′	6	2012 July 18	Summer
Bothnian Sea (Gävle)*	BÄH	100 (3)	N60°44′	E17°35′	6	2012 September 4	Autumn
Bothnian Sea (Hästskär)*	BH	50 (8)	N60°35′	E17°48′	6	2013 May 22	Spring
Central Baltic Sea (Vaxholm)*	BV	50	N59°26′	E18°18′	6	1979 August 27	Spring
Central Baltic Sea (Gamleby)*	BG	49	N57°50′	E16°27′	7	1979 August 20	Spring
Central Baltic Sea (Kalmar)*	BR	100	N57°39′	E17°07′	7	2012 May 09	Spring
Central Baltic Sea (Karlskrona)*	BA	100	N56°10′	E15°33′	7	2012 May 30	Spring
Central Baltic Sea*	BC	100	N55°24′	E15°51′	8	2011 October 18	Unknown
Southern Baltic Sea (Fehmarn)*	BF	50 (3)	N54°50′	E11°30′	12	1979 September 23	Autumn
Kattegat, Skagerrak, North Sea, East Atlantic Ocean							
Kattegat (Träslövs.läge)*	KT	50	N57°03′	E12°11′	20	1978 October 23	Unknown
Kattegat (Björköfjorden)*	KB	100	N57°43′	E11°42′	23	2012 March 12	Spring
Skagerrak (Brofjorden)*	SB	100	N58°19′	E11°21′	25	2012 March 20	Spring
Skagerrak (Hamburgsund)*	SH	49	N58°30′	E11°13′	25	1979 March 19	Spring
North Sea*	NS	49 (3)	N58°06′	E06°10′	35	1979 August 05	Autumn
Atlantic Ocean (Bergen)*	AB1	49	N64°52′	E10°15′	35	1980 February 07	Spring
Atlantic Ocean (Bergen)*	AB2	(8)	N60°35′	E05°00′	33	2013 May 22	Spring
Atlantic Ocean (Höfn)*	AI	100	N65°49′	W12°58′	35	2011 September 15	Spring
Pacific Ocean							
Strait of Georgia (Vancouver)*	PH	50 (6)	—	—	35	2012 November 24	—
West Atlantic Ocean							
Bonavista Bay†	WBoB	49 (2)	N48°49′	W53°20′	35	2014 June 25	Autumn
Fortune Bay†	WFB	50 (2)	N47°17′	W55°38′	35	2014 May 26	Spring
Inner Baie Des Chaleurs†	WInB	41 (2)	N48°00′	W65°51′	35	2014 May 8	Spring
Northumberland Strait†	WNsS	49 (2)	N46°19′	W64°09′	35	2014 May 6	Spring
Northumberland Strait†	WNsF	50 (2)	N45°44′	W62°36′	35	2014 September 16	Autumn
German Banks†	WGeB	48 (2)	N43°16′	W66°18′	35	2014 August 28	Autumn

Places where the sample was landed (if known) are given in parentheses. *n*, number of fish; *n2*, number of fish used for individual sequencing.

*Samples from our previous studies (11, 12).

†Samples sequenced in this study.

‡The spawning season for population samples has been predicted based on the stage of sexual maturation at the time of sampling.

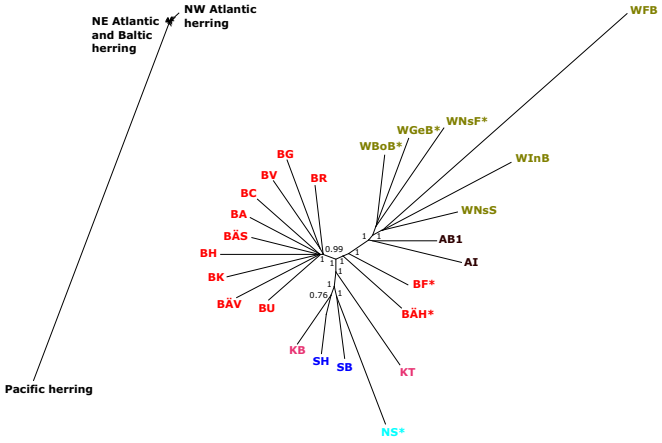

Fig. 2. (*Left*) Neighbor-joining phylogenetic tree. (*Right*) Zoom-in on the cluster of NE and NW Atlantic herring populations. Color codes for sampling locations are as in Fig. 1. Autumn-spawning populations are marked with an asterisk. To restrict the sampling variance and sequencing bias across the NE and NW populations, the phylogenetic tree was based on ∼1.2 million SNPs, each represented by 40–45 reads per population (*Materials and Methods*).

to 0.019 if we exclude the WFB population. This result is consistent with the tight clustering of populations in the phylogenetic tree (Fig. 2). This minute level of genetic differentiation is remarkable, given that our samples now include herring populations from both sides of the Atlantic, from the North Sea, Skagerrak, Kattegat, and the brackish Baltic Sea (Fig. 1). We performed a computer simulation study to investigate if the genetic differentiation among these 25 populations seemed to be driven primarily by selection or drift. We used 50,000 randomly sampled SNPs with 40–45 reads in each population and estimated F_{ST} (23) among all 25 populations for each locus. We then used simulation to generate a dataset reflecting the expected distribution of F_{ST} values for 50,000 selectively neutral loci based on 25 populations, each with an effective population size of N_e = 5,000 that were separated for t = 263 generations and have an expected F_{ST} identical to the one observed (F_{ST} = 0.026; when the WFB outlier population was excluded, the corresponding values were F_{ST} = 0.019 and t = 192 generations). The observed data deviated significantly from the simulated data because of a long tail of F_{ST} values >0.10 (Fig. 3). We conclude that the observed distribution deviates significantly from the one expected for selectively neutral alleles under a drift model, and the most plausible explanation is that this deviation is caused by natural selection.

Genetic Differentiation Between Spring- and Autumn-Spawning Populations. We explored the genomic regions showing differentiation between spring and autumn spawners from the NW Atlantic by comparing allele frequencies between the two groups SNP by SNP and identified 6,333 SNPs with significant differences in allele frequency ($P < 1 \times 10^{-10}$) (Fig. 4*A*, *Left*). These SNPs are conservatively estimated to represent at least 182 independent genomic loci (11). We compared these results with the loci that were associated with spawning time in the Baltic Sea/NE Atlantic populations (Fig. 4*A*, *Right*) (11). About 25% (1,747 of 6,333) of the associated SNPs in the NW Atlantic also reached statistical significance in the Baltic Sea/NE Atlantic comparison. The genetic signals associated with seasonal reproduction in the NE and NW Atlantic populations showed a considerable overlap (Fig. 4*B*). Six of eight previously identified genomic regions exhibiting the most consistent association with spawning time in NE Atlantic and Baltic populations replicated in the data from NW Atlantic (Dataset S1). At these six loci the same haplotype group is associated with autumn or spring spawning in all populations included in this study.

The results suggest that genetic factors affecting the timing of reproduction in herring are shared to a large extent among herring populations from both sides of the Atlantic. In contrast, genome-wide data indicate a closer genetic relationship between spring- and autumn-spawning herring from the same geographic region than between either spring- or autumn-spawning populations from different regions (Fig. 2, *Right*). Although many loci associated with the onset of reproduction were shared between the NE and NW Atlantic herring populations, there were certain genomic regions unique to each geographic region (Fig. 4 *A* and *B*). Apart from the timing of reproduction, spring- and autumn-spawning populations need to adapt to a variety of other ecological factors not considered in this study. The differentiated loci between spring and autumn spawners not shared between NE and NW Atlantic populations most likely reflect such local adaptations. The list of loci showing differentiation between spring and autumn spawners, the genes in their vicinity, and additional annotations of these regions are in Dataset S2.

Evidence of Parallel Evolution at *TSHR*, a Major Locus Associated with Timing of Reproduction. *TSHR*, encoding thyroid-stimulating hormone receptor, has a key role in photoperiodic regulation of reproduction in birds and mammals (24–26). The *TSHR* region showed the most significant difference in allele frequency between spring and autumn spawners in both the NE and NW Atlantic populations (Fig. 4*A*). The signatures of genetic differentiation within the ∼120-kb block around *TSHR* were strikingly similar in the NE and NW Atlantic populations (Fig. 4*C*). All autumn- and spring-spawning populations clustered separately, regardless of their geographic origin, in a neighbor-joining tree based on allele frequency data on 940 SNPs from this ∼120-kb region (Fig. 4*D*).

To reveal the haplotype structure at the *TSHR* locus, we used our whole-genome resequencing data of individual fish (16 spring-spawning and 9 autumn-spawning herring from NE Atlantic populations, 6 spring-spawning and 6 autumn-spawning herring from NW Atlantic populations, and 6 Pacific herring individuals as an outgroup). These data were used to generate a maximum likelihood tree for haplotypes at the 120-kb *TSHR* region (Fig. S3). Consistent with the results from pooled DNA sequencing (Fig. 4*D*), haplotypes from spring- and autumn-spawning individuals clustered as two distinct groups, and there was no clear substructuring related to geographic origin. The short branch lengths for *TSHR* haplotypes in spring spawners most likely reflect a recent selective sweep.

Two SNPs (upstream of *TSHR*) were found to be the most differentiated in the spring vs. autumn spawning contrast in both the NE and NW Atlantic herring populations (Fig. 4*B*). We generated phylogenetic conservation scores around these sites by comparing nine fish genomes, including Atlantic herring from our previous study (11). One of these SNPs (Scaffold1420: 133,030 bp) overlapped a conserved region. These highly differentiated *TSHR* SNPs are candidate mutations that may regulate *TSHR* expression in cells that are critical for the initiation of reproduction.

Haplotype Sharing Between Spring-Spawning and Autumn-Spawning Oceanic Herring Populations. We compared allele frequencies across all populations of herring at the loci showing highly significant differences in allele frequency between spring and autumn spawners, partly to explore whether the same alleles were associated with spawning time across the North Atlantic Ocean. The heat maps summarizing these data show two distinct patterns that we designate "A" and "B" (Fig. 5*A*). For loci belonging to pattern A, including the *TSHR* locus, differences in allele frequency between spring- and autumn-spawning herring

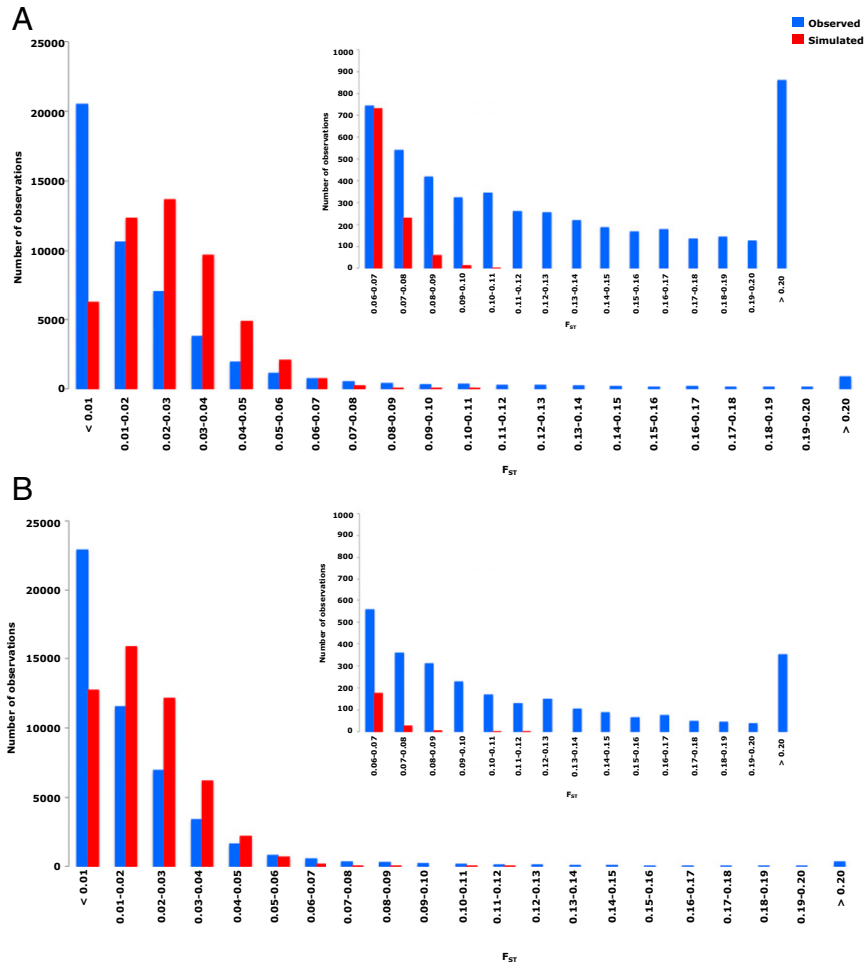

Fig. 3. Comparison of genetic differentiation among herring populations vs. the differentiation expected under genetic drift. The observed distribution is based on 50,000 randomly selected SNPs, and the simulated data are based on the same number of selectively neutral loci. Histograms show F_{ST} values in the simulated and observed datasets among all 25 Atlantic herring populations from both sides of the Atlantic Ocean (*A*) and in 24 populations excluding the WFB population (*B*). The right tail of the distribution is highlighted in the insets.

were remarkably consistent, independent of geographic origin. In contrast, in pattern B, involving SNPs located on eight different genomic scaffolds, the two oceanic spring-spawning herring populations sampled along the coasts of Norway (AB1) and Iceland (AI) were fixed for the same alleles as the autumn spawners from the NW and NE Atlantic, whereas the alternative alleles dominated in the majority of spring-spawning populations from both sides of the Atlantic (including the Baltic Sea) (Fig. 5*A*). Six of these eight genomic scaffolds include members of the myosin heavy chain (*MYHC*) gene family. Intermediate allele frequencies at the majority of these loci were observed in the two samples from Skagerrak (SB and SH), the transition area between the North Sea and the Baltic Sea (Figs. 1 and 5*A*).

A comparison with Pacific herring indicated that autumn spawners carried the ancestral allele at the great majority of SNPs (70.1%) belonging to pattern A, whereas they were associated with the derived allele at the majority of loci (69.9%) belonging to pattern B (Fig. 5*A*). The difference in the proportion of derived alleles associated with autumn spawning in patterns A and B is highly significant ($P = 2.0 \times 10^{-18}$, Fisher's exact test) and is most likely explained by linkage drag when alleles under selection have increased in frequency.

We further investigated individual haplotypes at the eight scaffolds associated with pattern B (Fig. 5*B*) in 43 fish, including spring and autumn spawners from both sides of the Atlantic and six Pacific herring (Table 1). These eight loci showed very strong linkage disequilibrium across populations even though they are spread in several genomic regions. All autumn-spawning herring were essentially fixed for the same allele associated with spawning season, whereas more heterogeneity was detected in spring-spawning populations (Fig. 5*B*). For example, considerable heterogeneity was observed in a spring-spawning population from the Baltic Sea (BH). If this heterogeneity represented sequence polymorphisms in a randomly mating population, one would expect that heterozygosity at these 189 SNPs from eight different scaffolds would be distributed more or less randomly among individuals, but in the BH population two individuals were heterozygous at most positions, and five individuals were essentially homozygous at all positions. Thus, a possible explanation is that the heterozygous fish represent hybrids between populations fixed for different alleles at these genomic regions. Similarly, in a spring-spawning population from the NW Atlantic (WFB) (Table 1), one fish was mostly homozygous for the autumn-spawning alleles, and the other was heterozygous at most diagnostic sites. The latter fish also was heterozygous at the *TSHR* locus (Fig. S3). The four individuals from the other two spring-spawning populations from the NW Atlantic (WInB and WNsS) were all homozygous for spring-spawning alleles (Fig. 5*B*). The eight fish sampled in spring 2013 close to Bergen, Norway (AB2) showed considerable heterogeneity at these eight loci, in contrast to the homogeneity observed in the sample AB1 collected in the same geographic area in February 1980 (Fig. 5*A*). Some fish were homozygous either for alleles associated with spring spawning or for alleles abundant in autumn-spawning and oceanic herring, whereas others had mixed haplotypes. It is possible that the AB2 sample represents a mix of individuals coming from different spring-spawning populations

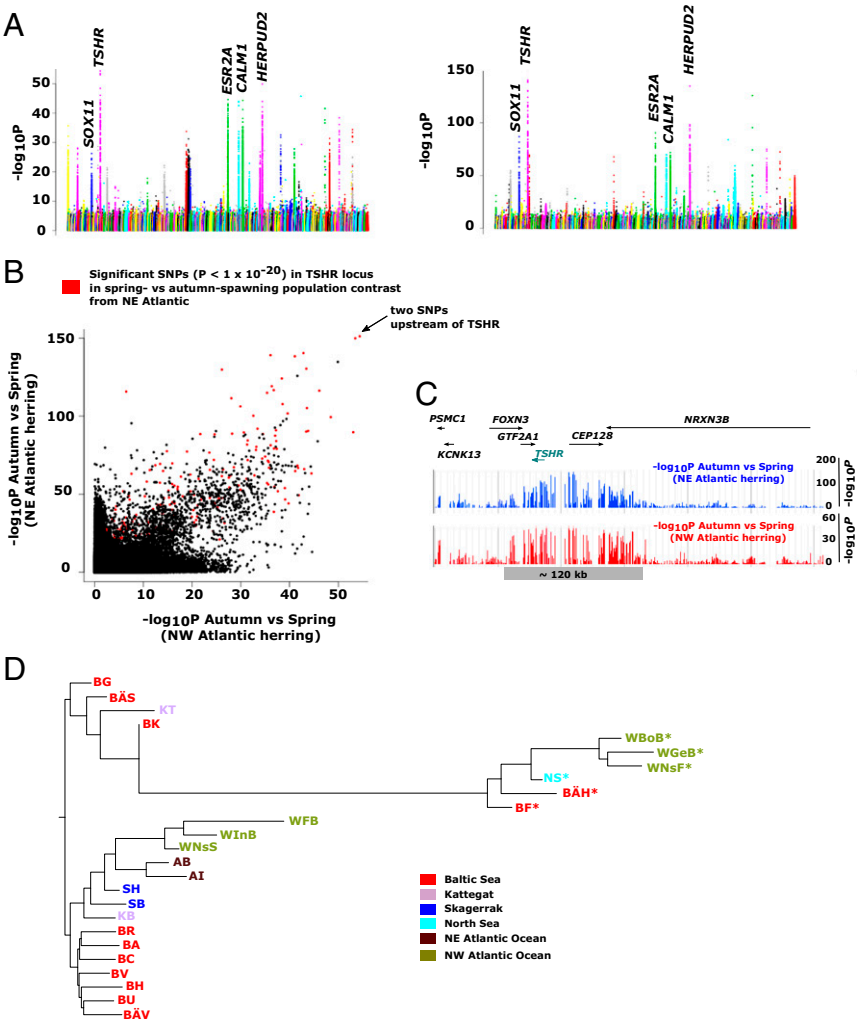

Fig. 4. Genetic differentiation between autumn- and spring-spawning herring. (*A*) Manhattan plot of *P* values for differences in allele frequency between autumn and spring spawners in the NW Atlantic herring (*Left*) and the NE Atlantic/Baltic herring (*Right*). (*B*) Correlation plot of *P* values for differences in allele frequency between autumn and spring spawners in NW Atlantic and in NE Atlantic/Baltic herring. (*C*) Distribution of *P* values around the 120-kb region harboring the *TSHR* locus. (*D*) Neighbor-joining tree based on allele frequencies, determined by pooled sequencing, for all SNPs (n = 1,313) in the *TSHR* region. Color codes for sampling locations are as in Fig. 1.

in Norway and their hybrids. Interestingly, data on somatic growth patterns and morphometric measurements indicated that interbreeding has occurred in the last 50 y between a resident, coastal spring-spawning population and the migratory, oceanic Norwegian spring-spawning herring (27); the sample AB1 in the current study is expected to represent this latter stock (Fig. 5*A*). This interbreeding likely took place after the collapse of the oceanic Norwegian spring-spawning population caused by overfishing in the late 1960s.

Validation of Loci Associated with Spawning Season. The genome-wide scans used for the detection of genetic differentiation among herring populations were based on the comparison of allele frequencies estimated from pooled DNA-sequencing data. Such data can be prone to biases caused by various factors, such as sequencing and mapping errors or differences in library or sequencing protocols, that could lead to uneven read depth distribution among loci and populations which could affect allele frequency estimates (28). Hence, to validate the results from pooled DNA sequencing, we genotyped a subset of highly differentiated SNPs between spring- and autumn-spawning populations in 377 individual fish (about 30 individuals from each of 13 populations). A total of 103 SNPs were selected using the following criteria: (*i*) those showing a consistent association with spawning time across the Atlantic and (*ii*) those unique to the NE or NW Atlantic. The individual genotyping results were consistent with the genetic signatures detected by pooled sequencing (Fig. S1).

For SNPs shared between the NE and NW Atlantic populations, the spring spawners tended to carry one haplotype and the autumn spawners tended to carry the alternative haplotype regardless of geographical location (Fig. S1*A*). For SNPs uniquely associated with spawning time in the NE Atlantic, there were two clear patterns (Fig. S1*B*). In the first pattern, the autumn spawners from the North Sea carried unique haplotypes not seen in any other population. In the second pattern, spring and summer spawners from the locations Gävle (BÄV and BÄS) and Kalix (BK) in the Baltic Sea (Table 1) shared a unique haplotype. For SNPs unique to the NW Atlantic, there was also a set of SNPs showing consistent differences between spring- and autumn-spawning populations only within the NW Atlantic (Fig. S1*C*). Some individuals from the autumn-spawning North Sea (NS) population segregated for some of these alleles present in the autumn-spawning populations in the NW Atlantic, whereas autumn-spawning herring from the Baltic Sea as well as all spring-spawning herring from the NE and NW Atlantic tended to be fixed for the alternative alleles.

Validation of Loci Associated with Adaptation to Low Salinity in the Brackish Baltic Sea. In our previous study (11), we identified about 3,000 SNPs representing about 100 independent genomic regions showing the most consistent correlation of allele frequencies with salinity by comparing populations adapted to the brackish

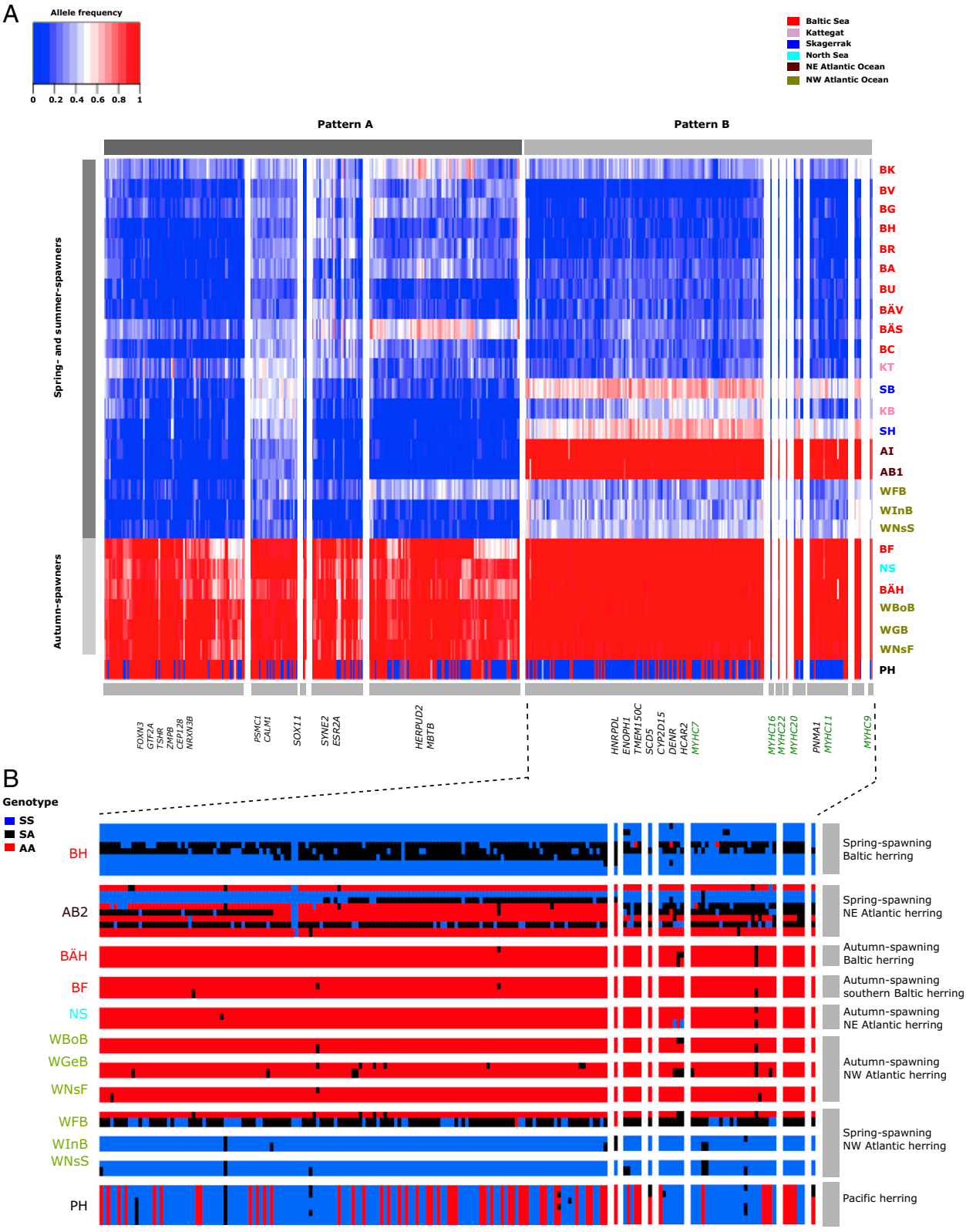

Fig. 5. Loci showing highly significant differences in allele frequency between spring- and autumn-spawning herring. (*A*) Heat map showing allele frequencies (estimated by pooled sequencing) in each of the 26 populations; color codes for each population are as in Fig. 1; genes overlapping these loci are listed at the bottom. (*B*) Haplotypes from individual herring samples at eight scaffolds showing pattern B (haplotype sharing between spring-spawning oceanic herring and autumn-spawning populations). The total number of SNPs used in these analyses is 189. Blocks of SNPs from different scaffolds are separated by blank lines.

Baltic Sea and the marine NE Atlantic (Dataset S3). The loci that are directly related to adaptations to low salinity are also expected to show strong genetic differentiation between the populations from the Baltic Sea and the NW Atlantic but not between the NE and NW Atlantic, where salinity is the same (35‰). To test this possibility, we calculated delta allele frequencies (dAF) between pairwise comparisons of Baltic, NW Atlantic, and NE Atlantic populations. The results showed that as many as 94.4% of the loci that showed a dAF >0.2 between the Baltic and NE Atlantic populations also showed a dAF >0.2 in the comparison of the Baltic vs. NW Atlantic populations (Fig. S2*A*). In contrast, few of these loci (12.7%) showed a dAF >0.2 in the contrast between the NE and NW Atlantic populations (Fig. S2*B*). Thus, the great majority of loci showing a consistent association with differences in salinity reported in our previous study were supported using the data from the NW Atlantic populations. The statistical support for differences in allele frequency in different pairwise comparisons among these groups is given in Dataset S3.

We previously identified a copy number variation (CNV) overlapping the gene for high choriolytic enzyme (*HCE*) that correlated with salinity. Populations from the brackish Baltic Sea (3–12‰) had a high copy number, whereas populations from marine waters in the NE Atlantic (20–35‰) had a relatively low copy number. As expected, all six NW Atlantic populations (35‰) showed a low copy number at this locus (Fig. S2*C*).

Discussion

Independent populations that are exposed to similar environmental conditions often evolve similar phenotypic traits. There are widespread examples of such parallel evolution in nature, and in some cases the genetic basis is known to some extent (e.g., refs. 5, 8, and 29). The Atlantic herring provides an opportunity to study the genetic basis of such repeated parallel evolution in geographically distant populations. In this study, we have demonstrated that genetic variants associated with adaptation to different spawning times (spring and autumn) are shared to a large extent among geographically distant populations of herring. This finding resembles the reuse of standing genetic variation for adaptation to marine and freshwater environments in the threespine stickleback (5). Thus, even though autumn-spawning herring from the Baltic Sea, the North Sea, and the NW Atlantic show genetic differentiation related to other traits (e.g., salinity) they share very similar haplotypes at loci that are strong candidates for underlying the timing of reproduction. This haplotype sharing implies that these variants were present in a common ancestor of these subpopulations or that they have been spread among populations via gene flow. The different populations are too closely related to allow us to distinguish efficiently which of these scenarios has been more important, but the very low levels of genetic differentiation among the 25 population samples included in this study ($F_{ST} = 0.026$) suggests a lack of reproductive isolation throughout the species range. This result is in sharp contrast to the situation for the Atlantic salmon, which shows strong genetic differentiation between populations from different continents (30, 31).

We identified a total of six independent loci that show a consistent association with spawning time across populations from the NW and NE Atlantic Ocean as well as in the brackish Baltic Sea (Fig. 4*A* and Dataset S1). It is likely that these loci contribute to how the timing of reproduction is determined, because one of the main environmental cues for the timing of reproduction, the change in day length, should be the same on both sides of the Atlantic. In contrast, the loci that are associated with spawning time only in one geographic region may reflect local adaptation. However, at present, the allele substitution effects at these six loci and the extent to which spawning time is genetically determined are unknown.

The *TSHR* locus shows the most convincing association with spawning time, because the two SNPs located upstream of this gene exhibit the most significant difference in allele frequency on the two sides of the Atlantic (Fig. 4*B*). We propose that one of these, if not both, is a causative change. Another interesting finding is the large haplotype block around *TSHR*, about 120 kb in size (Fig. 4*C*), associated with spawning time in contrast to the very rapid decay of linkage disequilibrium in parts of the herring genome not associated with ecological adaptation (11). The large haplotype block may be maintained by an inversion or, more likely, by the presence of multiple causative mutations across the associated region. In our previous study (11) we found no indications for the presence of an inversion at the *TSHR* locus, but the ability to detect inversions using short insert, paired reads is limited. The branch lengths in the phylogenetic tree for *TSHR* haplotypes associated with spring spawning are much shorter than the branch lengths for the haplotypes associated with autumn spawning, implying a more recent coalescence time for the former ones (Fig. S3). These data indicate that a relatively recent selective sweep has occurred in this region and that spring spawning may be a derived trait in the herring. There is, in fact, a clear trend that spring-spawning haplotypes also carry the derived allele at many of the SNPs (70.1%) showing the most consistent association with spawning time (Fig. 5*A*, pattern A); the derived state was deduced using the Pacific herring as an outgroup. Furthermore, phylogenetic trees generated from individual haplotypes for other loci showing strong association with spawning time also showed a general trend for a shorter coalescence time for spring-spawning haplotypes (Fig. S3). Thus, we propose that autumn spawning is the ancestral state in Atlantic herring.

Three of the loci showing the most consistent association with spawning time in herring contain genes (*TSHR*, *SOX11*, and *CALM*) with an established role in photoperiodic regulation of reproduction in birds and mammals (24–26, 32, 33). Functional studies are now required to confirm that these candidate loci also contribute to photoperiodic regulation of reproduction in herring. The robust associations reported here provide a unique opportunity to dissect the underlying molecular mechanisms for how increasing (spring) or decreasing (autumn) day length is translated to the initiation of spawning in different populations. The estrogen receptor beta 2 locus (*ESR2A*) has never been implicated in photoperiodic regulation of reproduction but has an established role in reproductive biology (34). Other loci that also have very convincing associations with spawning season, such as *HERPUD2* (homocysteine-responsive endoplasmic reticulum-resident ubiquitin-like domain member 2) and *SYNE2* (spectrin repeat containing nuclear envelope protein 2), have no known role in reproduction. Further studies on these associations provide an opportunity to establish new functional roles for these genes. *SYNE2* is, in fact, the neighboring gene to *ESR2* in all vertebrates sequenced to date, and the peaks of association in the two genes are only about 100 kb apart but are clearly separated by a region of weaker association (Fig. S4). It is an open question whether the *SYNE2* association is related to *SYNE2* function or if the region harbors long-range regulatory elements controlling *ESR2* expression (or vice versa). An association study such as this one reveals the location of causal mutations, but if these mutations are regulatory, the target gene(s) showing altered expression may be located outside the associated region.

One of the most interesting observations in this study was the unexpected haplotype sharing between spring-spawning and autumn-spawning oceanic populations. Members of the *MYHC* gene family were overrepresented in these regions. There are 23 annotated *MYHC* genes in the current herring genome assembly, and as many as six (26%) of these are located in these regions that constitute only 0.04% of the assembly. *MYHC* genes play a critical role for myogenesis (35). Previous studies have indicated that herring populations spawning at various times of

the year have a variable degree of developmental plasticity as regards myogenesis (36) and that differences in water temperature between spawning seasons are considered responsible for differential myogenesis in herring (37). Further research is required to reveal the ecological adaptation underlying the observed association.

This study has important practical implications for herring fishery in the NW Atlantic because it provides genetic markers that can distinguish spring- and autumn-spawning herring outside the breeding season. Such a diagnostic method could be applied to develop a more sustainable fishery by optimizing the intensity of fishing among stocks according to their abundance.

Materials and Methods

Sample Collection and DNA Extraction. Total genomic DNA was extracted from muscle tissue of 287 maturing individuals or individuals in spawning condition collected during the spawning season in 2014 at six localities on the east coast of Canada (NW Atlantic), from Newfoundland to the Scotian Shelf (Fig. 1 and Table 1). Tissue samples were stored in 95% ethanol at −20 °C until DNA isolation was performed following a standard phenol:chloroform:isoamyl alcohol protocol. DNA concentrations were measured as nanograms per microliter using the Quant-iT PicoGreen dsDNA Assay (Thermo Fisher Scientific) in a Roche LightCycler 480 Instrument. The integrity of DNA samples was verified with agarose gel electrophoresis and a molecular ladder in which nondegraded genomic DNA corresponds to a high-molecular-weight band around 23 kb.

Whole-Genome Resequencing and Variant Calling. Genomic DNA of 41–50 individuals per population were pooled in equimolar concentrations, resulting in six DNA pools, one per population. An aliquot of each pool was used for independent library preparations using the TruSeq Nano Illumina kit (Illumina). Fragment size selection was performed, in accordance with the instructions of the manufacturer using AMPURE beads, for an insert size between 450–550 bp. Each population library was paired-end sequenced at 126 cycles in one lane of a HiSeq-2500 sequencer to obtain a depth of coverage of >30× per pool. The quality of the raw reads was evaluated using FastQC v.10.1 (Brabraham Bioinformatics) (38), and low-quality reads were trimmed using Trimmomatic v.0.33 (39). The trimming of reads followed a sliding-window approach by which a read was cut at the 3′ end when the average PHRED33 quality score fell below 20 within a 5-bp window. Remaining Illumina adapter sequences were removed using the function ILLUMINACLIP (settings 2:30:10) implemented in Trimmomatic. Only pairs in which both reads were recovered after the quality-trimming step were used for downstream analysis.

The high-quality trimmed reads were aligned to the reference herring genome assembly (11) using default parameters of the algorithm BWA-MEM (v.0.6.2) (40). The sequences of the six NW Atlantic populations, together with the data of 19 Baltic Sea/NE Atlantic populations and one Pacific herring population from our previous study (11) were used to call SNP variants across all 26 populations. We followed the recommended workflows of the GATK tool (41, 42) for variant discovery. The raw variant calls were filtered using a stringent in-house filtering pipeline set up in our previous study (11). Various standard quality scores generated by GATK, such as SNP quality, mapping quality, base quality, mapping quality rank sum, read positions rank sum, allele frequency, and minimum and maximum read depth, were used to set up filtering parameters according to GATK best-practices recommendations (43). The cut-offs of these quality scores were chosen based on their genome-wide distributions. In addition, because the sequence data of the NE and NW Atlantic pools were generated at two different places (the NE Atlantic pool in Uppsala and the NW Atlantic pools in Halifax, Canada), there was a possibility of inflated genetic differentiation between the NE and NW Atlantic because of sequencing platform-specific technical bias. Hence, we applied more stringent hard filtering of these SNP quality scores, particularly while analyzing data for phylogenetic and NE vs. NW comparisons. Because individual sequencing for all NE and NW samples was done in Uppsala, we also used these individual data to evaluate the SNP calls from pooled DNA to exclude false calls that were specific to a sequencing platform.

To explore the haplotype structure at highly differentiated genomic regions, we performed further whole-genome resequencing of 27 individuals that included 12 samples from the NW Atlantic (two from each population used for pooled sequencing) and nine autumn-spawning samples from the NE Atlantic (Table 1). In addition, we also sequenced six individual Pacific herring samples to be used as outgroup (Table 1). Sequencing libraries were constructed with average fragment size of 400 bp, and 2 × 150-bp paired-end reads were generated using the Illumina HiSeq 2500 sequencing platform. Each library was sequenced to an ∼10× depth of coverage per individual. We combined these data with the whole-genome resequencing data of 16 spring-spawning NE populations (Table 1) from our previous study (11). The quality trimming, sequence alignments, and variant calling were done using a pipeline similar to that used for pooled sequencing described above.

Genome-Wide Screens for Genetic Differentiation. We separately combined the read depth count per SNP from the pooled sequencing data of three spring-spawning and three autumn-spawning NW Atlantic populations, obtaining two separate superpools, one for spring and one for autumn. Then we compared the allele frequency differences SNP by SNP between the superpools using a 2×2 contingency χ^2 test. We compared these results with the similar comparisons of autumn vs. spring spawners from the NE Atlantic populations of our previous study (11). In addition to the contingency χ^2 tests, we also screened for genetic differentiation using commonly used methods such as F_{ST} and pooled heterozygosity (44). The SNPs showing highly significant differentiation were clustered into independent genomic regions as described (11).

Simulations of Genetic Drift. The simulations aimed at assessing the expected distribution of F_{ST} under selective neutrality were conducted as in our previous paper on divergence among herring populations in the NE Atlantic (12). In brief, we used a slightly modified version of the Powsim software (45) that mimics sampling from populations at a predefined level of expected divergence through random number simulations under a classical Wright–Fisher model without migration or mutation. An infinitely large base population segregating for a specified number of independent, selectively neutral loci is divided into s subpopulations of equal effective size (N_e) through random sampling of $2N_e$ genes. Each of the subpopulations is allowed to drift for t generations, and the expected degree of divergence in generation t is then $F_{ST} = 1 - (1 - 1/2N_e)^t$ (46). The populations are sampled when the expected degree of differentiation has been reached; then F_{ST} (23) is calculated for each locus, and the distribution of simulated F_{ST} values is compared with the observed one.

To reduce statistical noise in the observed distribution, we restricted the analysis to SNPs that had 40–45 reads in all populations and calculated the F_{ST} for a random sample of 50,000 of these SNPs. The corresponding simulation was run with effective sizes of $N_e = 5{,}000$, the number of loci (SNPs) was 50,000, and the time of divergence (t) was set to result in an expected F_{ST} identical to the average of that found in the observed distribution.

Phylogenetic Analysis. Genetic divergence between populations was calculated using PLINK (47), and phylogenetic trees based on allele frequencies estimated from pooled sequencing data were generated using PHYLIP (48). Genetic distances were calculated using an identity by state (IBS) similarity matrix (Dataset S4). The bootstrapping of the phylogenetic tree was done using the Phylip Seqboot package (48). The phased haplotypes were generated from the genomic regions showing high differentiation among populations using BEAGLE (49), and maximum likelihood haplotype trees were built generated using FastTree (50).

CNV Analysis of the *HCE* Locus. We used GATK:DepthOfCoverage to extract the read depth coverage of all populations. All reads with mapping quality below 20 were filtered out. We then generated 1-kb nonoverlapping windows and normalized them against the AB1 pool that had the highest sequence coverage among all samples in our previous study (11). We scanned the genome to identify CNV between autumn and spring spawners from both side of Atlantic by selecting the following populations:

i) Autumn spawners: WBoB, WGeB, WNsF, BF, BAH, and NS
ii) Spring spawners: WFB, WInB, WNsS, BAV, BH, and AB1

After filtering windows with low depth, we compared both groups in 715,093 windows by ANOVA.

Genotyping Individual Fish in a Subset of SNPs. To validate the candidate loci, we genotyped 384 individuals (192 from the NW Atlantic and 192 from the NE Atlantic populations), using a Sequenom SNP panel. To do so, we chose 105 SNPs with the following three criteria: (*i*) 35 SNPs showing highly significant differences in allele frequency between autumn- and spring-spawning populations that are shared in herring from the NW and NE Atlantic populations; (*ii*) 35 SNPs that are unique to herring from the NE Atlantic and show highly significant differences in allele frequency between autumn- and spring-spawning populations; and (*iii*) 35 SNPs that are unique to herring from the NW Atlantic and show highly significant differences in allele frequency between autumn- and spring-spawning populations.

The genotype data were analyzed for standard quality control using PLINK (47). SNPs with missing genotypes in >10% of the individuals were excluded. Similarly, individuals that had missing genotypes in >10% of SNPs were excluded. One hundred three SNPs genotyped in 377 individuals passed these thresholds and were used for the subsequent downstream analysis. The allele frequency estimates from individual genotyping were compared with the estimates from pooled whole-genome sequencing. The genotyping results were consistent with the results based on pooled sequencing. The correlation coefficients of allele frequency estimates from individual and pooled sequencing were in the range $r = 0.95–0.97$ in respective populations. The haplotype structures of the candidate loci based on Sequenom genotyping of individual fish resembled those generated by pooled sequencing (Fig. 4 and Fig. S2).

ACKNOWLEDGMENTS. This work was funded by the European Research Council project BATESON (L.A.), the GENSINC project funded by the Norwegian Research Council (L.A.), and Discovery Grant RGPIN-2014-04696 (to D.E.R.). A.P.F.-P. was funded by a Vanier Canada Graduate Scholarship, the Killam Trust, and a Nova Scotia Graduate Scholarship. Simulation work by N. Ryman was supported by the Swedish Research Council and the Swedish Research Council Formas grants to Linda Laikre. Pooled DNA sequencing of NW Atlantic populations was performed at The Centre for Applied Genomics of the Hospital for Sick Children, Toronto. Sequencing of individual samples was performed by the SNP&SEQ Technology Platform, which is supported by Uppsala University and Hospital, the Science for Life Laboratory, and Swedish Research Council Grants 80576801 and 70374401. Computer resources were provided by the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX), Uppsala University.

1. Miller MR, et al. (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Mol Ecol* 21:237–249.
2. Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nat Rev Genet* 14:807–820.
3. Pearse DE, Pogson GH (2000) Parallel evolution of the melanic form of the California legless lizard, Anniella pulchra, inferred from mitochondrial DNA sequence variation. *Evolution* 54:1041–1046.
4. Hoekstra HE, Nachman MW (2003) Different genes underlie adaptive melanism in different populations of rock pocket mice. *Mol Ecol* 12:1185–1194.
5. Jones FC, et al.; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.
6. Stern DL (2013) The genetic causes of convergent evolution. *Nat Rev Genet* 14: 751–764.
7. Pearse DE, Miller MR, Abadia-Cardoso A, Garza JC (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proc Biol Sci* 281:20140012.
8. Hohenlohe PA, et al. (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet* 6:e1000862.
9. Pritchard JK, Di Rienzo A (2010) Adaptation - not by sweeps alone. *Nat Rev Genet* 11:665–667.
10. Luo Y, Widmer A, Karrenberg S (2015) The roles of genetic drift and natural selection in quantitative trait divergence along an altitudinal gradient in Arabidopsis thaliana. *Heredity (Edinb)* 114:220–228.
11. Martinez Barrio A, et al. (2016) The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *eLife* 5:e12081.
12. Lamichhaney S, et al. (2012) Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring. *Proc Natl Acad Sci USA* 109: 19345–19350.
13. McQuinn I (1997) Metapopulations and the Atlantic herring. *Rev Fish Biol Fish* 7: 297–329.
14. Iles TD, Sinclair M (1982) Atlantic herring: Stock discreteness and abundance. *Science* 215:627–633.
15. Gröger JP, Kruse GH, Rohlf N (2010) Slave to the rhythm: How large-scale climate cycles trigger herring (Clupea harengus) regeneration in the North Sea. *ICES J Mar Sci* 67:454–465.
16. Sinclair M, Tremblay MJ (1984) Timing of spawning of Atlantic herring (Clupea harengus harengus) populations and the match–mismatch theory. *Can J Fish Aquat Sci* 41:1055–1065.
17. Department of Fisheries and Oceans Canada (2011) Canadian fisheries statistics 2008. Available online at www.dfo-mpo.gc.ca/stats/commercial/cfs/2008/CFS2008_e.pdf. Accessed August 1, 2016.
18. Stephenson RL, Melvin GD, Power MJ (2009) Population integrity and connectivity in Northwest Atlantic herring: A review of assumptions and evidence. *ICES J Mar Sci* 66: 1733–1739.
19. Reid RN, et al. (1999) Atlantic herring, Clupea harengus, life history and habitat characteristics. Available online at www.nefsc.noaa.gov/publications/tm/tm126/tm126.pdf. Accessed August 1, 2016.
20. Department of Fisheries & Oceans Canada (2012) Assessment of Atlantic herring in the southern Gulf of St. Lawrence (NAFO Div. 4T). Available online at www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2012/2012_014-eng.pdf. Accessed August 1, 2016.
21. McPherson A, O'Reilly P, Taggart C (2004) Genetic differentiation, temporal stability, and the absence of isolation by distance among Atlantic Herring populations. *Trans Am Fish Soc* 133:434–446.
22. McPherson AA, Stephenson RL, O'Reilly PT, Jones MW, Taggart CT (2001) Genetic diversity of coastal northwest Atlantic herring populations: Implications for management. *J Fish Biol* 59:356–370.
23. Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution (N Y)* 38:1358–1370.
24. Ono H, et al. (2008) Involvement of thyrotropin in photoperiodic signal transduction in mice. *Proc Natl Acad Sci USA* 105:18238–18242.
25. Hanon EA, et al. (2008) Ancestral TSH mechanism signals summer in a photoperiodic mammal. *Curr Biol* 18:1147–1152.
26. Nakao N, et al. (2008) Thyrotrophin in the pars tuberalis triggers photoperiodic response. *Nature* 452:317–322.
27. Johannessen A, et al. (2014) The dynamics of a metapopulation: Changes in life-history traits in resident herring that co-occur with oceanic herring during spawning. *PLoS One* 9:e102462.
28. Schlötterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals - mining genome-wide polymorphism data without big funding. *Nat Rev Genet* 15: 749–763.
29. Bradbury IR, et al. (2010) Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proc R Soc London* 277:3725–3734.
30. Stahl G (1983) Differences in the amount and distribution of genetic variation between natural populations and hatchery stocks of Atlantic salmon. *Aquaculture* 33: 23–32.
31. King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski BA (2001) Population structure of Atlantic salmon (Salmo salar L.): A range-wide perspective from microsatellite DNA variation. *Mol Ecol* 10:807–821.
32. Kim H-D, et al. (2011) Class-C SOX transcription factors control GnRH gene expression via the intronic transcriptional enhancer. *Mol Endocrinol* 25:1184–1196.
33. Melamed P, et al. (2012) Gonadotrophin-releasing hormone signalling downstream of calmodulin. *J Neuroendocrinol* 24:1463–1475.
34. Bondesson M, Hao R, Lin C-Y, Williams C, Gustafsson J-Å (2015) Estrogen receptor signaling during vertebrate development. *Biochim Biophys Acta* 1849:142–151.
35. Watabe S (1999) Myogenic regulatory factors and muscle differentiation during ontogeny in fish. *J Fish Biol* 55:1–18.
36. Johnston IA, Vieira VLA, Temple GK (2001) Functional consequences and population differences in the developmental plasticity of muscle to temperature in Atlantic herring Clupea harengus. *Mar Ecol Prog Ser* 213:285–300.
37. Temple GK, Cole NJ, Johnston IA (2001) Embryonic temperature and the relative timing of muscle-specific genes during development in herring (Clupea harengus L.). *J Exp Biol* 204:3629–3637.
38. Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. Available online at www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed August 1, 2016.
39. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
40. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.
41. Van der Auwera GA, et al. (2013) From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma* 43: 11.10.1–33.
42. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303.
43. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498.
44. Rubin C-J, et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464:587–591.
45. Ryman N, Palm S (2006) POWSIM: A computer program for assessing statistical power when testing for genetic differentiation. *Mol Ecol Notes* 6:600–602.
46. Nei M (1987) *Molecular Evolutionary Genetics* (Columbia Univ Press, New York).
47. Purcell S, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575.
48. Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5: 164–166.
49. Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81:1084–1097.
50. Price MN, Dehal PS, Arkin AP (2010) FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490.