Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring

Sangeet Lamichhaney[a,1], Alvaro Martinez Barrio[a,1], Nima Rafati[a,1], Görel Sundström[a,1], Carl-Johan Rubin[a], Elizabeth R. Gilbert[a,2], Jonas Berglund[a], Anna Wetterbom[b], Linda Laikre[c], Matthew T. Webster[a], Manfred Grabherr[a], Nils Ryman[c], and Leif Andersson[a,d,3]

[a]Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden; [b]Science for Life Laboratory, Department of Cell and Molecular Biology, Karolinska Institutet, SE-17177 Stockholm, Sweden; [c]Department of Zoology, Stockholm University, SE-10691 Stockholm, Sweden; and [d]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75124 Uppsala, Sweden

Contributed by Leif Andersson, September 26, 2012 (sent for review July 28, 2012)

The Atlantic herring (*Clupea harengus*), one of the most abundant marine fishes in the world, has historically been a critical food source in Northern Europe. It is one of the few marine species that can reproduce throughout the brackish salinity gradient of the Baltic Sea. Previous studies based on few genetic markers have revealed a conspicuous lack of genetic differentiation between geographic regions, consistent with huge population sizes and minute genetic drift. Here, we present a cost-effective genome-wide study in a species that lacks a genome sequence. We first assembled a muscle transcriptome and then aligned genomic reads to the transcripts, creating an "exome assembly," capturing both exons and flanking sequences. We then resequenced pools of fish from a wide geographic range, including the Northeast Atlantic, as well as different regions in the Baltic Sea, aligned the reads to the exome assembly, and identified 440,817 SNPs. The great majority of SNPs showed no appreciable differences in allele frequency among populations; however, several thousand SNPs showed striking differences, some approaching fixation for different alleles. The contrast between low genetic differentiation at most loci and striking differences at others implies that the latter category primarily reflects natural selection. A simulation study confirmed that the distribution of the fixation index F_{ST} deviated significantly from expectation for selectively neutral loci. This study provides insights concerning the population structure of an important marine fish and establishes the Atlantic herring as a model for population genetic studies of adaptation and natural selection.

Baltic herring | genetics | population biology

The Atlantic herring is a pelagic fish that constitutes an enormous biomass in the North Atlantic waters. The number of fish in a school of Atlantic herring may exceed 1 billion. It has a key role in the North Atlantic ecosystem and constitutes the major prey for many marine mammals, birds, squid, and other fishes. Herring has been a critical food resource in Northern Europe, and its major utility at present is for producing fish feed for the aquaculture industry, in addition to direct human consumption. It is one of the few marine species that is able to reproduce in both the saline Atlantic and throughout the brackish Baltic Sea (Fig. 1*A*). It, thus, shows adaptation to a well-defined environmental variable (salinity) that must have taken place within the last 16,000 y, subsequent to the last glaciation (Fig. 1*B*). Since medieval times, the herring in the Gulf of Bothnia and Central Baltic Sea has been named "Baltic herring" (in Swedish "strömming"), based on its distinct phenotype (e.g., smaller size and lower fat content compared with the herring in the Atlantic). Linnaeus (1) classified the Baltic herring as a subspecies, *Clupea harengus membras* (L.), of the Atlantic herring (*Clupea harengus*, L.). Genetic adaptation in the herring may also occur because of differences in spawning time, temperature (2), light conditions (3), feed resources, predators, or other variables. For instance, the Atlantic herring is an important prey for whales in the Atlantic Ocean, predators that are not present in the Baltic Sea.

An accurate description of the population structure of the Atlantic herring is crucial for sustainable utilization of this important species. However, the degree of genetic differentiation among stocks is largely unknown. Early studies based on a dozen isozyme loci revealed conspicuously low levels of genetic differentiation between geographically distant and morphologically distinct forms (4, 5). More recent studies using microsatellites and a limited number of SNPs confirmed this pattern, but one microsatellite and 16 SNPs showed significant genetic differentiation between regions consistent with selection acting at a subset of loci (2, 6–8). Here, we expand these analyses by analyzing patterns of genetic differentiation on a genome-wide scale and identifying genetic markers likely to be under selection that control phenotypic variation.

Ideally, a genome-wide analysis requires access to a high-quality draft genome; however, the construction of such an assembly for a large genome like the herring, on the order of 900 Mbp (9), is still a major undertaking. Whereas it is easy to generate sufficient sequence coverage of the genome through high-throughput technologies, constructing the long jumping libraries required to bridge highly repetitive regions, and to correctly order contigs along the chromosome, remains a technical challenge. Here, we applied a strategy that does not require a high-quality draft genome: combining a transcriptome assembly with whole-genome shotgun sequencing allows for constructing an "exome assembly" (i.e., exons obtained from assembled mRNA transcripts with flanking sequences derived from genomic reads) and then perform a genome-wide screen for genetic polymorphisms. This approach enabled a comprehensive analysis that allowed us to make a major advance with regard to the genetic characterization of the Atlantic herring. We identified over 440,000 SNPs and found strong genetic differentiation at a small fraction of these (about 2–3%), which can only be explained by natural selection.

Results

Transcriptome Analysis and Exome Assembly. Skeletal muscle mRNA was isolated from a single Baltic herring caught in the
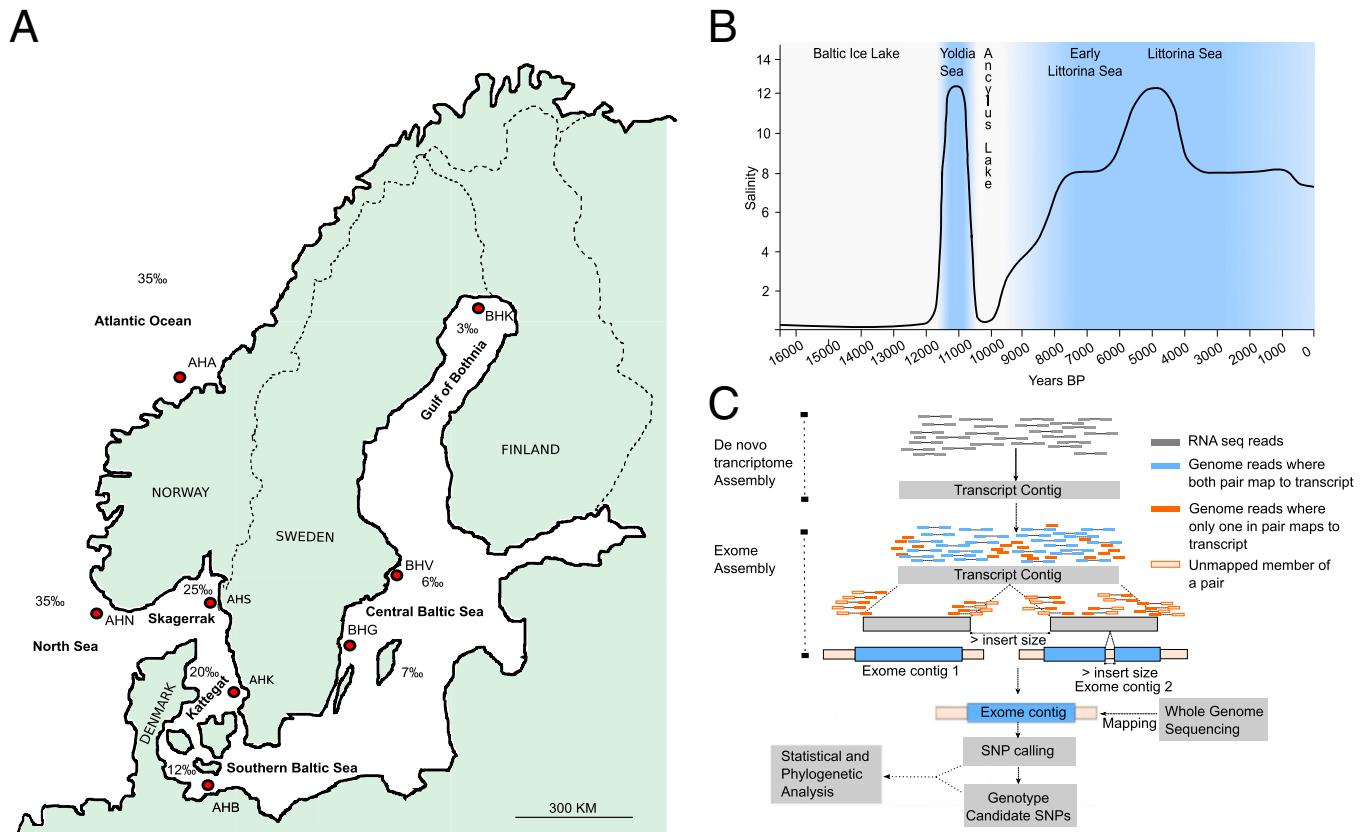
GENETICS

Fig. 1. Description of sample collection and experimental design. (*A*) Map with sample localities and the gradient of salinity, from 3‰ in the inner Gulf of Bothnia to >30‰ in the Atlantic Ocean. The abbreviations of sample localities are given in Table 1. (*B*) The development of the Baltic Sea is usually divided in four different periods. The earliest is the Baltic Ice Lake (16,000–11,600 B.P.), a freshwater basin, followed by the Yoldia Sea (11,600–10,700 B.P.) that includes a brackish phase. Subsequently, the freshwater Ancylus Lake (10,700–8,500 B.P.) occurred, followed by the Littorina Sea (8,500 B.P. to present) when the Baltic Sea turns brackish again (20). High-resolution strontium isotope analyses of mollusc shells indicate that the salinity in the first phase of the Littorina Sea (7,130–2,775 B.P.) was around 12‰ compared with ~7‰ today (21). (*C*) Construction of the exome assembly and SNP calling. First, RNA seq reads were assembled into transcript contigs using Trinity (10). Second, genomic reads from one population were aligned to the transcript contigs to generate exome contigs. Third, SNP calling was performed after aligning genomic reads from pooled samples to the exome assembly.

archipelago of Stockholm (Sweden). We generated a total of 116 million reads of 101 bp in size and de novo assembled these data into a transcriptome using Trinity (10). We then aligned genomic reads from one of the samples described below (Gulf of Bothnia) to the transcripts, thus extending the transcriptome into an exome assembly. This strategy allows for identifying intron/exon boundaries ([Fig. S1](#)) and also provides better alignability of genomic reads at exon/intron

Table 1. Information about samples of Atlantic and Baltic herring collected during the period 1978-1980

Locality*	Sample	N	E	Salinity (‰)	Date (yy/mm/dd)	Spawning season	$P^{\dagger}$	$H^{\ddagger}$
Baltic herring								
Gulf of Bothnia (Kalix)	BHK	65°52′	22°43′	3	800629	Spring	84.8	22.6
Central Baltic Sea (Vaxholm)	BHV	59°26′	18°18′	6	790827	Spring	83.3	22.7
Central Baltic Sea (Gamleby)	BHG	57°50′	16°27′	7	790820	Spring	85.5	23.1
Atlantic herring								
Southern Baltic Sea (Fehmarn)	AHB	54°50′	11°30′	12	790923	Autumn	89.0	23.2
Kattegat (Träslövsläge)	AHK	57°03′	12°11′	20	781023	Spring	84.1	22.9
Skagerrak (Hamburgsund)	AHS	58°30′	11°13′	25	790319	Spring	86.6	23.0
North Sea	AHN	58°06′	06°10′	35	790805	Autumn	87.0	22.9
Atlantic Ocean	AHA	64°52′	10°15′	35	800207	Winter-spring	86.6	22.8

  AHA, Atlantic herring Atlantic Ocean; AHB, Atlantic herring Southern Baltic Sea; AHK, Atlantic herring Kattegat; AHN, Atlantic herring North Sea; AHS, Atlantic herring Skagerrak; BHG, Baltic herring Gamleby; BHK, Baltic herring Kalix; BHV, Baltic herring Vaxholm.
*Places where the sample was landed (if known) are given in parenthesis.
†Proportion of the total SNPs that have read support for both alleles in each sample.
‡Average heterozygosity = sum of [2·*p*·(1 − *p*)] for all SNPs ÷ total number of SNPs, where *p* is the frequency of the most common allele.
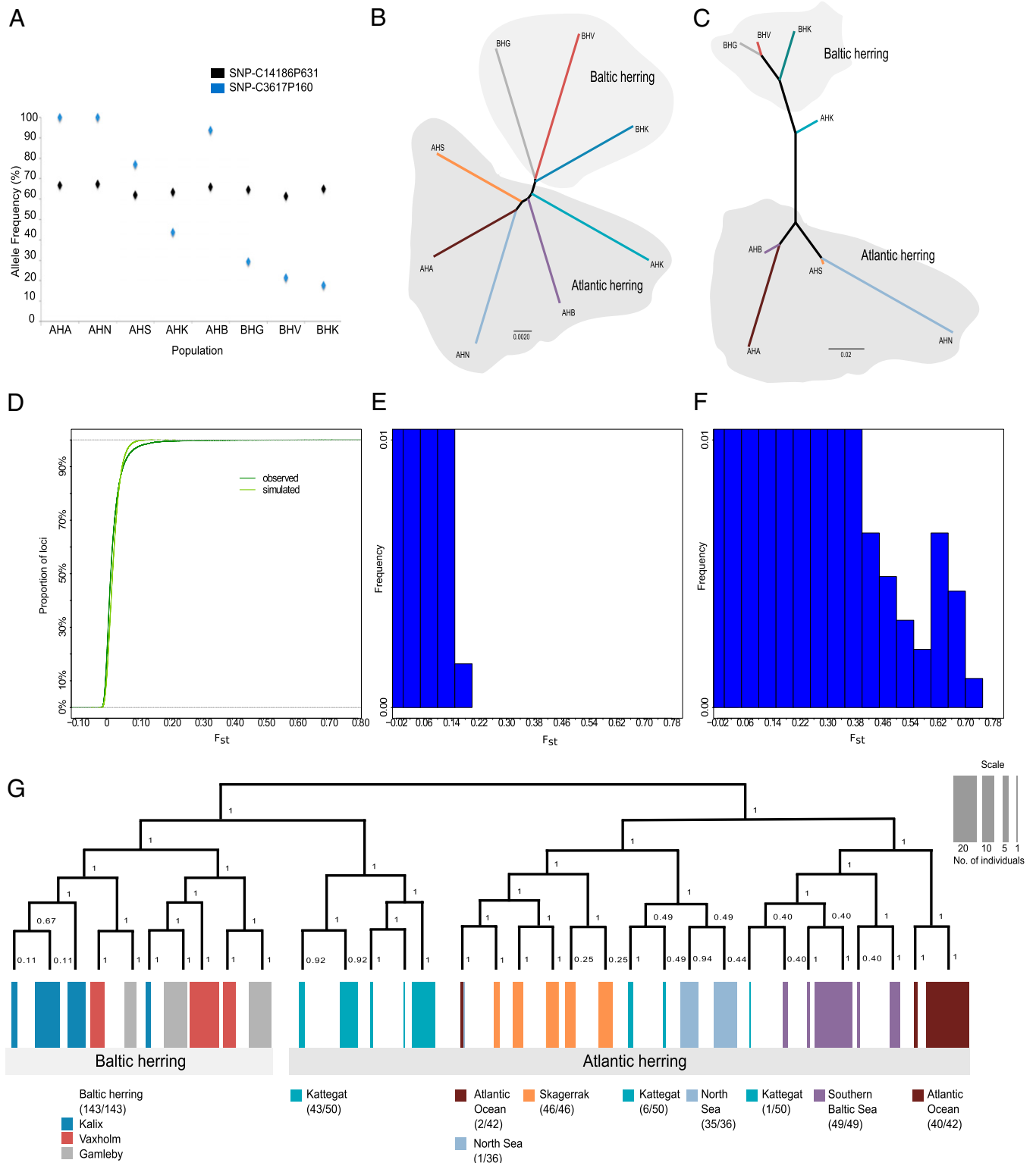
Fig. 2. Analysis of genetic relationships and simulation analysis of the F_{ST} distribution. (*A*) Examples of SNPs showing no significant (C14186P631) and highly significant (C3617P160) genetic differentiation among localities. (*B* and *C*) Phylogenetic tree analysis based on all markers (*B*) and based on those showing highly significant genetic differentiation (*C*). (*D*) Cumulative distribution of observed and simulated (assuming neutrality) F_{ST} values. (*E* and *F*) Histogram of F_{ST} values in the simulated (*E*) and observed (*F*) dataset (note the truncated *y* axis). (*G*) Clustering of individual herring using fineSTRUCTURE (11) based on individual SNP genotype data for 1,583 loci. The number of individual herring of each category placed in the same cluster is given in parenthesis. The size of the bars reflects the number of individuals in each cluster as indicated by the scale bar. Abbreviations for name of localities are given in Table 1.

borders in the SNP pipeline (Fig. 1*C*). The statistics of the transcriptome and exome assembly are summarized in Table S1.

Whole-Genome Sequencing of Pooled Samples. We sequenced eight pools of 50 fish each from different geographic locations that were sampled during 1979–1980 (5) (Fig. 1*A*), and each pool was

sequenced to about 30-fold genome coverage. We aligned the reads to the exome assembly and estimated allele frequencies for a total of 440,817 SNPs identified by our SNP calling pipeline (Fig. 1C). The proportion of polymorphic loci and the expected heterozygosities were very similar across populations (Table 1), consistent with low genetic differentiation between regions. For the great majority of loci, contingency χ^2 tests did not reveal any significant allele frequency differences among samples (Fig. 2A). A phylogenetic tree based on all SNPs indicated a star-like phylogeny, consistent with low genetic differentiation (Fig. 2B), although there was a weak phylogenetic signal in the tree with the three samples of Baltic herring clustered on one side of the tree and the Atlantic Ocean and North Sea samples at the other side. A small set of SNPs ($n = 3,847$) showed highly significant genetic differentiation ($P < 10^{-10}$; Table S2 and Fig. 2A), and a phylogenetic tree built from these SNPs shows clear separation of populations (Fig. 2C). We refer to these SNPs as "outlier" SNPs and those that did not show a significant genetic differentiation as "neutral"; the latter does not strictly mean that they are selectively neutral, only that they were associated with genomic regions that did not show genetic differentiation in this study.

Comparison of the Observed and a Simulated Distribution of F_{ST} Values. To investigate whether the observed genetic differentiation is primarily driven by drift or selection, we performed a simulation study to explore whether the distribution of the fixation index F_{ST} in our data are consistent with the expectation for selectively neutral loci. To reduce the sampling variance in F_{ST} values, we restricted this analysis to those 36,794 SNPs represented by at least 40 reads in each population. We then performed a simulation study based on eight subpopulations, each with an effective population size of 10,000 individuals that had been separated for 451 generations, resulting in an average F_{ST} identical to the one observed ($F_{ST} = 0.022$). There was, in general, good agreement between the observed and simulated data (Fig. 2D). About 90% of the loci, both in the simulated and observed data, had low F_{ST} values in the range 0–0.05. However,

there was a highly significant excess ($P \approx 0$) of loci with extreme F_{ST} values in the real data compared with the simulation (Fig. 2 E and F), indicating that selection is the dominant force causing genetic differentiation. For example, 200 loci in the observed data had $F_{ST} > 0.1883$, whereas none was found in the simulated data. Furthermore, 3.7% of the loci in the observed data ($n = 1,350$) had $F_{ST} > 0.0849$, whereas the corresponding value in the simulated data were 1% ($n = 368$). We conclude that 2–3% of the loci in our genome-wide screen show more genetic differentiation than is expected under a selectively neutral genetic drift model.

SNP Genotyping of Individual Fish. We validated the sequence-based allele frequency estimates by individual genotyping of the same fish as used for pooled sequencing. A total of 1,244 neutral and 1,583 outlier SNPs were successfully genotyped in the majority of the 400 individuals (Table S3). There was a strong correlation between the two types of allele frequency estimates (Fig. S2). A cluster analysis based on individual SNP data did not reveal any clustering at all for the neutral SNPs, whereas we observed striking clustering using the outlier SNPs, albeit not in perfect agreement with the tree based on allele frequency estimates from sequencing pooled DNA samples as expected (Fig. S3). We also used fineSTRUCTURE (11) for an unbiased search of genetic similarities among the 400 individuals included in this study. This software did not reveal any significant clustering of individuals when we used the 1,244 neutral SNPs. In sharp contrast, the outlier SNPs revealed a remarkable clustering of fish by proximity of location (Fig. 2G). All fish classified as Baltic herring formed one major group that clustered with Atlantic herring sampled from Kattegat, close to the straight in which saltwater from the Atlantic mixes with brackish water coming out of the Baltic Sea (Fig. 1A), whereas all other samples of Atlantic herring formed the other major group. A unique genetic status of the Atlantic herring population in the North Sea/Baltic Sea transition zone was recently noted also by Limborg et al. (2).
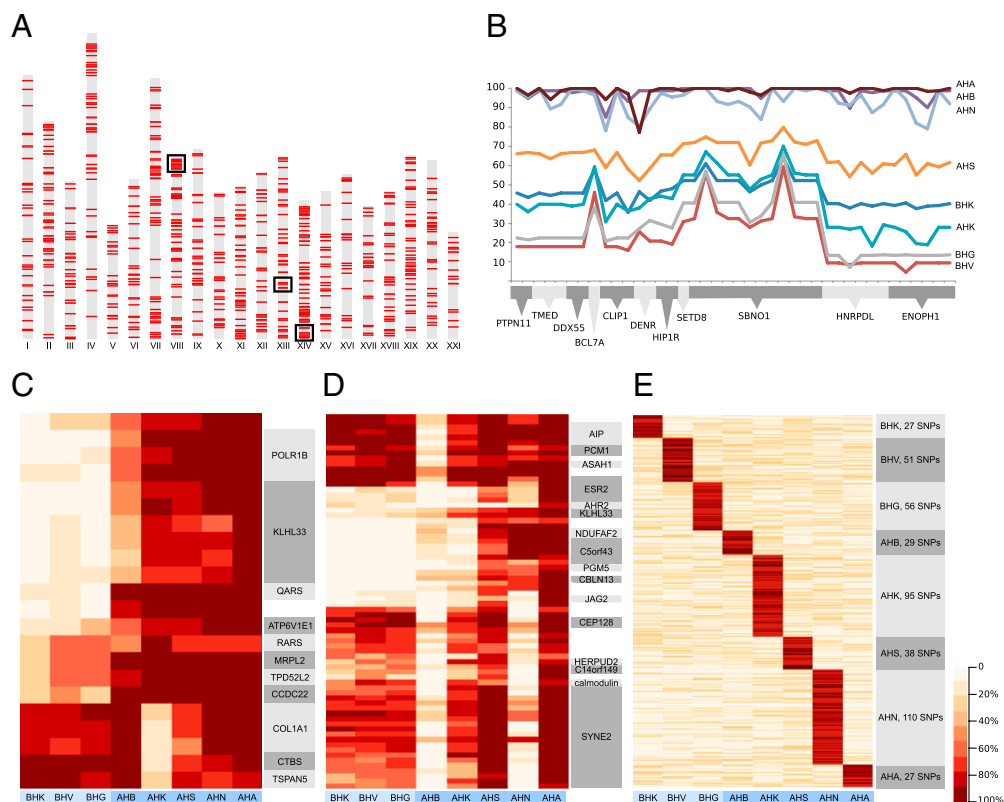

Fig. 3. Genomic distribution and allele frequencies at genetically differentiated loci. (A) Distribution of significant SNP loci plotted on the stickleback genome assembly. Three clusters of differentiated loci are marked by boxes. (B) Allele frequencies among samples at a cluster of 40 SNPs corresponding to a region on stickleback chromosome XIV. Transcript names are given below the x axis. (C–E) Heat maps illustrating striking allele frequency differences in different contrasts of sample locations. Samples of Baltic herring vs. Kattegat (C), autumn-spawning Atlantic herring from the Southern Baltic Sea vs. Atlantic Ocean (D) and those SNPs where the major allele in one population was the minor allele in all other populations (E). Designations of transcripts (based on BLAST hits) are given to the right of C and D. The most common allele in the sample of Atlantic herring (AHA) was used as the reference allele at all loci in B–D. Abbreviations for name of localities are given in Table 1.

Pattern of Genetic Differentiation Among Herring Populations. Having established that a small percentage of the loci detected in this study show more genetic differentiation than expected from genetic drift, we set out to explore the genome-wide distribution of the most significant loci ($P < 10^{-10}$); these included 3,847 SNPs with F_{ST} values in the range of 0.09–0.91. To investigate possible genomic colocation of the SNPs in the absence of a genome assembly for the herring, we gained a preliminary view by first extracting the transcripts associated with outlier SNPs and then mapping the transcripts onto the orthologs in the stickleback (*Gasterosteus aculeatus*) genome (12) (Fig. 3*A*), taking advantage of the high degree of conserved synteny among teleost fish (13). Loci under selection in the herring are spread across the stickleback genome but with some clustering. Three prominent clusters included 64 SNPs in a block containing 20 genes on stickleback chromosome VIII, 51 SNPs in 10 genes on stickleback chromosome XIII, and 98 SNPs in 14 genes on stickleback chromosome XIV. In particular, the loci in the chromosome XIV cluster showed a remarkable pattern of variation in which the allele frequencies were very similar across loci within populations. This pattern is even more striking when plotting the allele frequencies for SNPs that have been genotyped in individual fish (Fig. 3*B*). An examination of individual genotypes showed that the region is divided into three blocks in which the linkage disequilibrium (LD) among loci approaches 100% (Fig. S4*A*), explaining the similarity in allele frequencies across loci. The samples from Southern Baltic Sea (autumn-spawning Atlantic herring), North Sea, and the Atlantic Ocean were close to fixation for the same alleles at these loci, whereas these alleles constituted the minor allele at most loci in the three samples of Baltic herring and in the Kattegat (the straight between the Atlantic and the Baltic) sample. By contrast, the sample from Skagerrak, the region in the Atlantic adjacent to Kattegat (Fig. 1), had intermediate allele frequencies at all loci, consistent with its geographic location in between the other populations that show more marked differences. The allele frequency distributions in this cluster of 98 SNPs are in perfect agreement with the overall picture revealed by fineSTRUCTURE, with most Atlantic herring from the Kattegat sample clustering with the Baltic herring (Fig. 2*G*). Furthermore, the homozygosity across the region in the samples from the Atlantic Ocean, North Sea, and Southern Baltic Sea implies strong positive selection for one favored haplotype (Fig. S4*A*). The two clusters on chromosomes VIII and XIII were also composed of haplotype blocks with strong LD (Fig. S4 *B* and *C*). The data imply that several clusters of loci, most likely maintained by suppressed recombination, have contributed to genetic differentiation in the herring. The presence of inversions is one plausible mechanism for suppression of recombination.

Because salinity in the Kattegat, the entrance to the Baltic Sea, can change dramatically with saltwater flowing in from the Atlantic and brackish water flowing out from the Baltic Sea, we next examined our data for loci distinguishing the Kattegat sample (AHK), from Baltic herring (BHK) despite their similarity in allele frequencies at most loci. Twenty-two loci in 11 genes stood out in this contrast ($P < 10^{-10}$). The allele frequencies at many of these loci correlated well with salinity at the geographic locations of the samples (compare Figs. 1*A* and 3*C*). Indeed, one of the genes showing this pattern was *ATP6V1E1* (*ATPase, H⁺ transporting, lysosomal 31kDa, V1 subunit E1*), which encodes a subunit of a vacuolar H⁺-ATPase proton pump involved in osmoregulation (14). This locus could, therefore, be important for fine-scale adaptation to local levels of salinity.

In terms of spawning season, 72 loci showed marked allele frequency differences between autumn-spawning Atlantic herring caught in the Southern Baltic (AHB) and winter-spawning herring from Atlantic Ocean (AHA) (Fig. 3*D*), despite similar allele frequency estimates at most other SNPs, including the ones that are significantly different between Atlantic and Baltic herring. We can divide these differentiated loci into three types: (*i*) the major allele in Southern Baltic was the minor allele in all other samples; (*ii*) the Southern Baltic fish clustered with other samples from the brackish Baltic Sea; and (*iii*) Southern Baltic and North Sea shared similar allele frequency estimates but differed from all other populations. Interestingly, Southern Baltic and North Sea represent the only autumn-spawning herrings included in this study, whereas all other populations are winter- or spring-spawners, suggesting that one or more of these loci may influence reproduction. Finally, we searched our data for loci where the major allele in an individual sample was the minor allele in all other samples (Fig. 3*E*). This showed that all populations had "private" SNP alleles fulfilling this criteria and the number of such SNPs were in the range of 27–110, with the Atlantic herring samples from Kattegat and North Sea having the largest number of private alleles.

Discussion

Overall, this study has confirmed low levels of genetic differentiation at neutral loci in the herring, a finding consistent with earlier reports based on few loci. It is a breakthrough with regard to describing the population structure of the Atlantic herring and provides compelling evidence for the existence of a number of genetically differentiated populations of herring in the North Atlantic. The results support Linnaeus (1) classification of the Baltic herring as a subspecies. The genetic differences among the three samples of Baltic herring are small compared with the allele frequency differences between Baltic and Atlantic herring, except for one sample of Atlantic herring that was caught at the border between Kattegat and the Baltic Sea (Fig. 1*A*). This sample is surprisingly similar to Baltic herring populations at many, but not all, loci showing genetic differentiation between Baltic herrings and other populations of Atlantic herring. This may indicate that this sample either represents a population that is more closely related to an ancestor of the Baltic herring populations or represents a population that breeds in the brackish Baltic Sea but migrates into Kattegat for feeding. The Kattegat sample was classified as a spring-spawning herring collected outside the spawning season (Table 1). The autumn-spawning herring from the Southern Baltic Sea were more similar to Atlantic herring from Skagerrak, the North Sea, and the Atlantic Ocean than to Baltic herring samples, consistent with its classification as an Atlantic herring.

The number of adult herring in the Baltic Sea alone is on the order of 20 billion individuals. Furthermore, a single female can produce 20,000–60,000 eggs (15), and the herring spawns in huge shoals where eggs and sperm are mixed. Thus, herring populations are expected to have enormous effective population sizes. The low genetic differentiation at most loci is in agreement with large effective population sizes and minute genetic drift. Furthermore, there may also be some gene flow between subpopulations that contribute to the lack of genetic differentiation at selectively neutral loci. The large effective population sizes imply that natural selection is the dominating force that determines the frequency of nonneutral alleles. The results of this study, revealing low genetic differentiation at the majority of loci contrasted by marked genetic differentiation at a small percentage of the loci, is consistent with this view. If genetic differentiation is primarily determined by natural selection, the pattern of genetic similarities among populations will vary from locus to locus as illustrated in Fig. 3 *C–E*.

The herring is distributed in a salinity gradient from the inner parts of the Baltic Sea to the open Atlantic Ocean (Fig. 1*A*). Interestingly, a number of loci previously associated with response to salinity and osmoregulation in other species showed strong genetic differentiation that correlated with salinity. These include *ATP6V1E1* (see above) as well as another subunit of the same pump, *ATP6V0E1* (*ATPase, H⁺ transporting, lysosomal 9kDa, V0 subunit E1*). Significant SNPs were identified in *TMED2* (*transmembrane emp24 domain trafficking protein 2*), also known as *P24A*, which encodes a protein that controls the transport of a calcium sensing receptor whose expression is salinity dependent in tilapia (*Oreochromis mossambicus*) (16, 17). We also found strong genetic differentiation at the hemoglobin α (*HBA*)

and β (*HBB*) loci; hemoglobin β has been shown to be up-regulated in tilapia after exposure to saltwater (18). Interestingly, the allele frequency of a synonymous SNP in *HBA* was recently reported to be significantly correlated with salinity in Atlantic and Baltic herring (2).

This study opens up a plethora of possibilities for follow-up studies. For instance, how stable are allele frequencies at selected loci over time? The samples included in this study were collected during 1978–1980, and since then, there has been a clear trend with milder winters involving reduced extension and duration of ice cover in the Baltic Sea, changing the abiotic environment for herring. Furthermore, the average size of adult Baltic herring has halved in the last 30 y (19), probably because of the increased population size of sprat (*Sprattus sprattus*) that effectively compete with the herring for zooplankton as food source, but intense commercial fishing may also have contributed to this trend. Have these changes in abiotic and biotic conditions affected the Baltic herring population composition or allele frequencies at selected loci within just a few decades? Another interesting change that has occurred since the 1970s is that autumn-spawning herring have nearly disappeared in the Central Baltic Sea. Is this because previously autumn-spawning herring have changed their reproductive behavior or because they have been eradicated by overfishing? Our results have important implications for sustainable fishery management of the herring. We have revealed thousands of genetic markers that can be used for efficient monitoring of herring populations. The cost-effective approach developed in this study can easily be applied to any fish species exploited by commercial fishing, as well as to any organism for which there is a need for genome-wide analysis.

The present study is a major advance toward the genetic characterization of the Atlantic herring. However, the study has its limitations because it is based on an exome assembly built on a transcriptome representing a single tissue (skeletal muscle). The interesting results of this study justify further work, including a more comprehensive transcriptome study (more tissues), as well as the development of a high-quality draft genome assembly. The latter would allow us to define the number of independent loci that show genetic differentiation. In the present study, we were able to get some initial insight into this pattern by aligning our transcript contigs to the stickleback genome. This indicated that loci under selection show a genome-wide distribution but with some striking clustering. The presence of inversions is one possible explanation for the observation of such cluster of loci showing very strong linkage disequilibrium. Inversions were identified to have an important role for maintaining contrasting ecotypes of the three-spine stickleback adapted to marine and freshwater environments (12). The present study suggests that a similar mechanism may be operating in the Atlantic and Baltic herring with regard to its adaptation to different environmental conditions.

This study establishes the herring as a model organism for evolutionary genetics because of the potential to study the effects of natural selection in populations in which drift plays a subordinate role. Distinguishing drift from selection is a major challenge in population genetic studies, and there is much debate over the proportion of loci under selection, proportion of substitutions fixed by positive selection, the prevalence of hard vs. soft sweeps, and the distribution of fitness effects of new mutations. Many of these points of debate could be addressed by analyzing a model organism in which the effects of genetic drift can be excluded from the analysis.

Materials and Methods

The *Materials and Methods* are described in detail in *SI Materials and Methods*, which includes the following subjects: transcriptome analysis, whole-genome sequencing, exome assembly, SNP detection, statistical analysis, SNP genotyping, phylogeny analysis, and identifying patterns of strong selection.

1. Linnaeus C (1761) *Fauna Suecica* [*Sistens Animalia Sueciae regni ...*], 2nd Ed. Stockholm: Laurentius Salvius.
2. Limborg MT, et al.; FPT Consortium (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Mol Ecol* 21(15):3686–3703.
3. Jokela-Määttä M, Smura T, Aaltonen A, Ala-Laurila P, Donner K (2007) Visual pigments of Baltic Sea fishes of marine and limnic origin. *Vis Neurosci* 24(3):389–398.
4. Andersson L, Ryman N, Rosenberg R, Ståhl G (1981) Genetic variability in Atlantic herring (*Clupea harengus harengus*): Description of protein loci and population data. *Hereditas* 95(1):69–78.
5. Ryman N, Lagercrantz U, Andersson L, Chakraborty R, Rosenberg R (1984) Lack of correspondence between genetic and morphological variability patterns in Atlantic herring (*Clupea harengus*). *Heredity* 53(3):687–704.
6. Larsson LC, et al. (2007) Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. *Mol Ecol* 16(6):1135–1147.
7. Larsson LC, Laikre L, André C, Dahlgren TG, Ryman N (2010) Temporally stable genetic structure of heavily exploited Atlantic herring (*Clupea harengus*) in Swedish waters. *Heredity (Edinb)* 104(1):40–51.
8. Gaggiotti OE, et al. (2009) Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* 63(11):2939–2951.
9. Hardie DC, Hebert PDN (2004) Genome-size evolution in fishes. *Can J Fish Aquat Sci* 61(9):1636–1646.
10. Grabherr MG, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7):644–652.
11. Lawson DJ, Hellenthal G, Myers S, Falush D (2012) Inference of population structure using dense haplotype data. *PLoS Genet* 8(1):e1002453.
12. 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(7392):55–61.
13. Sarropoulou E, Fernandes JM (2011) Comparative genomics in teleost species: Knowledge transfer by linking the genomes of model and non-model fish species. *Comp Biochem Physiol Part D Genomics Proteomics* 6(1):92–102.
14. Guffey S, Esbaugh A, Grosell M (2011) Regulation of apical H+-ATPase activity and intestinal HCO3− secretion in marine fish osmoregulation. *Am J Physiol Regul Integr Comp Physiol* 301(6):R1682–R1691.
15. Kändler R, Dutt S (1958) Fecundity of Baltic herring. *ICES Rapports et Procès-Verbaux des Réunions* 143(2):99–108.
16. Loretz CA, Pollina C, Hyodo S, Takei Y (2009) Extracellular calcium-sensing receptor distribution in osmoregulatory and endocrine tissues of the tilapia. *Gen Comp Endocrinol* 161(2):216–228.
17. Stepanchick A, Breitwieser GE (2010) The cargo receptor p24A facilitates calcium sensing receptor maturation and stabilization in the early secretory pathway. *Biochem Biophys Res Commun* 395(1):136–140.
18. Rengmark AH, et al. (2007) Identification and mapping of genes associated with salt tolerance in tilapia. *J Fish Biol* 71(Suppl C):409–422.
19. Casini M, et al. (2010) Linking fisheries, trophic interactions and climate: Threshold dynamics drive herring *Clupea harengus* growth in the central Baltic Sea. *Mar Ecol Prog Ser* 413:241–252.
20. Andren T, et al. (2011) *The Development of the Baltic Sea Basin During the last 130 ka Central and Eastern European Development Studies*, eds Harff J, Bjorck S, Hoth P (Springer, Berlin), Vol 3, pp 75–97.
21. Widerlund A, Andersson PS (2011) Late Holocene freshening of the Baltic Sea derived from high-resolution strontium isotope analyses of mollusk shells. *Geology* 39(2):187–190.