

Supplementary Information Appendix

Sperm, egg and embryo proteins critical for genetic adaptation of herring to low salinity in the Baltic Sea

Cheng Ma, Fahime Mohamadnejad Sangdehi, Mari Kawaguchi, Kaori Sano, Svenja V. Da. Pettersson, Andreas Wallberg, Joshua L. Wort, Yumeng Yan, Sergei Moshkovskii, Florian Berg, Arild Folkvord, Christof Lenz, Henning Urlaub. U. Benjamin Kaupp, Shigeki Yasumasu and Leif Andersson

Material and Methods Supplementary Text

Supplementary Fig. 1. Genome-wide association analysis between Atlantic and Baltic herring populations. (A) Manhattan plot showing results obtained using logistic regression implemented in PLINK, treating Baltic individuals ($n = 20$) as cases and Atlantic individuals ($n = 75$) as controls. Genomic control correction was applied to adjust for potential residual population stratification. The y-axis represents $-\log_{10}(P\text{-values})$. The dashed red line indicates the Bonferroni-corrected genome-wide significance threshold ($P < 8.1 \times 10^{-9}$). Highlighted loci correspond to candidate genes located within the strongest association peaks. (B) Quantile–quantile (QQ) plot showing the distribution of observed versus expected $-\log_{10}(P\text{-values})$, used to assess overall calibration of test statistics and deviation from the null expectation.

Supplementary Fig. 2. Allele frequency heatmap showing highly differentiated SNPs at the *LRRC8C2* locus in a complete set of population samples from Pacific, Atlantic and Baltic herring. Purple boxes above the heatmap indicate missense positions. Note that *LRRC8C2* is encoded on the negative strand, reversing the SNP order.

Supplementary Fig. 3. PRM LC-MS/MS analysis of three of the four monitored *LRRC8C2* peptides in herring sperm. These *LRRC8C2* peptides were readily detected, whereas the *LRRC8C1* peptide LYIYNDGTK was not detected. The panels show exemplary PRM transition channels with retention time on the x-axis and MS signal intensity on the y-axis. (A) MS signal intensity of endogenous (upper panel) and heavy (lower panel) LGNNLLSGLSPK peptide corresponding to *LRRC8C2*, in herring sperm peptide background; 10 fmol of heavy peptide were loaded on column. (B) MS signal intensity of endogenous (upper panel) and heavy (lower panel) VHVEEGNLLYK peptide corresponding to *LRRC8C2* in herring sperm peptide background; 10 fmol of heavy peptide were loaded on the column. (C) MS signal intensity of endogenous (upper panel) and heavy (lower panel) LYFSHNK peptide corresponding to *LRRC8C2* in herring sperm peptide background; 10 fmol of heavy peptide were loaded on column. (D) MS signal intensity of endogenous (upper panel) and heavy (lower panel) LYIYNDGTK peptide corresponding to *LRRC8C1* in herring sperm peptide background; 0.005 fmol of heavy peptide were loaded on the column. Endogenous peptide was not detected.

Supplementary Fig. 4. PRM LC-MS/MS analysis of the *LRRC8C2* peptide VHVEEGNLLYK in liver, muscle, brain, and heart. Exemplary transition channels from the most sensitive measurements are shown, with retention time on the x-axis and MS signal intensity on the y-axis. Each panel shows MS signal intensities of endogenous (upper panel)

and heavy (lower panel) VHVEEGNLLYK peptide in the respective tissue peptide background. For all tissue samples, the lowest determined LOD was 0.005 fmol per µg of protein. Endogenous peptide was not detected in any tissue. (A) Liver. (B) Muscle. (C) Brain. (D) Heart.

Supplementary Fig. 5. Sequence alignment using ClustalX of the proteins LRRC8C2 (Chr2) Atlantic (A0A6P3VMJ8), LRRC8C2 Intermediate allele, LRRC8C2 Baltic North allele, LRRC8C2 Baltic South allele, and LRRC8C1 (Chr10) (A0A6P8G6J1). Single amino-acid substitution within the sequences of A0A6P3VMJ8, Intermediate allele, Baltic north, and Baltic south alleles are in red. Peptides of the proteins identified by mass spectrometry in the sperm sample are highlighted in yellow. Two identified, unique peptides for LRRC8C Baltic north allele, which are not present in the other isoforms, are highlighted in green. Four of the common peptides located in the 100% identical protein region also match to the protein LRRC8C1 (Chr10) (UniProt A0A6P8G6J1). No unique peptides for LRRC8C1 (Chr10) (A0A6P8G6J1) could be identified which indicates the absence of this isoform in sperm. For detailed information of identified LRRC8C peptides, see Dataset S1a.

Supplementary Fig. 6. Electrostatic maps of the pore opening of various LRRC8 channels. (A) Surface representation of the ESD of hexameric *H. sapiens* LRRC8A. Arrows indicate the outer ring of negative charges (at the channel's orifice) conferred by residue D100 (red) and the inner ring (within the pore lumen) of positive charge conferred by residues R103 and H104 (blue). (B) Surface representation of the ESD of heptameric *M. musculus* LRRC8C. (C) Surface representation of the ESD of heptameric *C. harengus* LRRC8C1. (D) Surface representation of the ESD of heptameric *C. harengus* LRRC8C2. All electrostatic maps were calculated using the APBS plug-in of Pymol. An arrow indicates the ring of negative charge conferred by residue D103. The charge density scale is indicated (and is the same for (A)-(C)).

Supplementary Fig. 7. Bird's-eye view of the selectivity filters of various LRRC8 channels. (A) Overlay of monomer structures of LRRC8: *M. musculus* C (PDB: 8B40), *H. sapiens* A (PDB: 5ZSU), *M. musculus* C and *H. sapiens* A (AlphaFold2), *C. harengus* C1 of chromosome 10 and C2 of chromosome 2 (AlphaFold2) shown in grey, cyan, green, magenta, blue, and orange cartoon representation. The residues corresponding to the single-nucleotide polymorphisms (SNPs): D25, V254, K452, and L515 in *C. harengus* *membras* are shown as yellow spheres. (B) Cartoon of the ESD of hexameric human LRRC8A channel. (C) Cartoon of the ESD of heptameric *M. musculus* LRRC8C channel. (D) Cartoon of the ESD of heptameric *C. harengus* LRRC8C1 subunit. (E) Cartoon of the ESD of heptameric *C. harengus* LRRC8C2 subunit. In all structures, subunits are indicated H1-6 or H1-7, clockwise from centre-top, and residues implicated in anion selectivity (A), or at homologous positions (B)-(E) are labelled, colour-coded, and shown in stick representation.

Supplementary Fig. 8. Pore profiles of various LRRC8 channels. (A) (left) Cartoon of hexameric *H. sapiens* LRRC8A (PDB: 5ZSU), with 4 of 6 subunits removed. (right) The pore profile. (B) (left) Cartoon of heptameric *M. musculus* LRRC8C (PDB: 8B40), with 5 of 7 subunits removed. (right) The pore radius profile. (C) Structure of heptameric *C. harengus* LRRC8C1 (AF2) shown in cartoon representation, with 5 of 7 subunits removed. (right) The pore profile. (D) Cartoon of heptameric *C. harengus* LRRC8C2 (AF2), with 5 of 7 subunits removed. In all structures the pores are shown as grey spheres, and residues at homologous positions to P15, T48, R103, and K235 in human LRRC8A are shown in stick representation. (right) The pore profile. All pore profiles were calculated using the MOLE online webserver (mole.upol.cz), and the approximate positions of residues at each constriction are labelled.

Residues missing from electron density maps are marked with an asterisk and the position was estimated by aligning an AF2 structure of monomeric LRRC8 with a subunit of each homomultimer. (E) An overlay of all pore profiles for comparison.

Supplementary Fig. 9. Allele frequency contrasts comparing Atlantic spring-spawners against Baltic spring-spawners for the region on chromosome 22 harbouring *ZPBA1*.

SNPs marked in red represent missense mutations.

Supplementary Fig. 10. Allele frequency heatmap showing highly differentiated SNPs at the *FTG* locus on chromosome 17 in a complete set of population samples of Pacific, Atlantic and Baltic herring. Purple boxes in the top row indicate missense positions.

Supplementary Fig. 11. Allele frequency heatmap showing highly differentiated SNPs at the *ZPBA1* locus in a complete set of population samples of Pacific, Atlantic and Baltic herring. Purple boxes in the top row indicate missense positions.

Supplementary Fig. S12. Individual haplotype neighbor-joining trees from the *FTG*, *ZPBA1*, and *HE1C* loci. (A) *FTG* (Chr 17:25.338 - 25.401 Mb); (B) *ZPBA1* (Chr 22:20.871 - 20.874 Mb); (C) *HE1C* (Chr 26:4.993 - 4.999 Mb). Branch lengths represent nucleotide divergence (i.e. average difference per base) across the included regions.

Supplementary Fig. 13. Clustering of individual haplotypes across the *FTG* locus. (A) Hierarchical clustering based on Hamming edit distance, where each edit corresponds to one nucleotide difference, between individual haplotypes across the *FTG* locus (chr 17: 25:34-25.40 Mb), showing that all Baltic haplotypes form a sub-clade inside the larger Pacific clade, without abnormal branch length. The Atlantic haplotypes found inside the Baltic clade stem from two coastal individuals caught at Askøy, Norway. (B) as (A), but based only on non-synonymous positions, showing that, on the amino acid level, there are Pacific haplotypes that are indistinguishable from those found in the Baltic. Also noteworthy is a single haplotype from Vancouver that is distinct from its peers, and instead cluster with the Subarctic/Baltic types.

Supplementary Fig. 14. Observation of isolated egg envelopes of Atlantic and Baltic herring in water droplets. Developing eggs of Atlantic and Baltic herring were crashed, and the egg envelopes were rinsed in PBS. The isolated egg envelopes of Atlantic (A) and Baltic herring (B) were placed in PBS droplets and observed under a binocular microscope. When water in the droplets was removed, the Atlantic egg envelopes collapsed (C), whereas the Baltic egg envelopes maintained a round shape (D), suggesting that Atlantic egg envelopes are softer than those of Baltic herring. Scale bar = 500 µm.

Supplementary Fig. 15. Observation of egg and egg envelope. Unfertilized (A–D) and fertilized (E–H) eggs of Baltic and Atlantic herring was compared. The 4% PFA fixed eggs were rehydrated and observed using a binocular microscope (A, B, E, F). HE-stained sections of the egg envelopes are shown in panels C, D, G, H. The lower side of the image corresponds to the cytoplasmic side. ad: adhesive layer, in: inner layer. Scale bars in (A, B, E, F): 1000 µm, and (C, D, G, H): 5 µm. The section of the egg envelopes was generated as follows. Fertilized and unfertilized eggs were fixed in 4% paraformaldehyde/PBS and stored in 70% ethanol at –30 °C until use. After stepwise replacement with PBS, the eggs were embedded in 5% agarose (Ultra-low Gelling Temperature Agarose, Sigma–Aldrich, MO, USA) prepared with 20% sucrose and then frozen. Frozen sections were cut at 14 µm

thickness. After removal of agarose, sections were placed in Mayer's Hematoxylin Solution (FUJIFILM Wako Chemical Co., Osaka, Japan) for 2 min, washed in running water for 15 min, and placed in 0.5% Eosin Y Solution (FUJIFILM Wako Chemical Co., Osaka, Japan) for 2 min. After dehydration with ethanol, they were placed in xylene for 15 min and sealed in LIMO mount (Pharma Co., Ltd., Tokyo, Japan).

Supplementary Fig. 16. Maximum-likelihood phylogeny of fish hatching enzyme (HE) sequences from Atlantic herring and other fish. The tree was inferred with a codon substitution model using sequences corresponding to mature hatching enzyme proteins. Tip labels in blue and green denote Atlantic herring and European sprat sequences generated in this study, respectively. Tip labels for herring and sprat sequences follow the format Genus_species_chr#_position (Mb). Clades highlighted in red, orange and purple correspond to *HE1*, *HE2* and *HE3* genes in Clupeidae, respectively. Node values indicate bootstrap support from 2,000 replicates; values below 50% are not shown. The scale bar indicates branch length.

Supplementary Fig. 17. Amino acid polymorphism in the mature *HE1* protein of Atlantic herring. Amino acid sequence alignment showing only polymorphic sites within the mature *HE1* coding region identified in the Atlantic herring reference genome (Ch_v3.1). Sequence identifiers indicate genomic locations (chr#_position, Mb). A sequence from chromosome 26 is shown as the reference (top row); residues identical to the reference are denoted by dots, substitutions by letters, and gaps by dashes. Numbers above and below the alignment indicate variable-site indices and corresponding residue positions, respectively.

Supplementary Fig. 18. Expression of *HE1C* genes in developing embryos of Atlantic and Baltic herring. mRNA expression measured as RPKM (reads per kilobase million) in developing embryos 6, 12, and 18 days after fertilization. The minute expression indicated for Atlantic herring is most likely misalignment of reads from other HE loci.

Supplementary Fig. 19. Partial purification of hatching enzyme. One hundred mL of hatching liquid (culture medium collected after embryo hatching) from Atlantic or Baltic herring were dialyzed against 25 mM Tris-HCl (pH 7.5) containing 0.05% Brij 35 (Polyoxyethylene 23 Lauryl Ether), and applied to a Toyopearl SP-650M cation-exchange column (Tosoh Corp., Tokyo, Japan) equilibrated with the same buffer. The column was washed with the same buffer, and the adsorbed proteins were eluted in a single step using 25 mM Tris-HCl (pH 7.5) containing 1 M NaCl and 0.05% Brij 35. Panels (A) and (B) show the elution profiles of hatching enzymes from Atlantic and Baltic herring, respectively, on Toyopearl SP-650M column chromatography. (C) Casein zymography patterns of the partially purified hatching enzyme of Atlantic herring (left) and Baltic herring (right). The numbers on the left represent the size of the molecular markers. Each zymogram yielded a single band at 24 kDa, which is in agreement with values calculated from mature enzyme sequences of cDNAs (*HE1*-chr26, MW 22,657.0 and *HE1*-chr22, MW 22,911.3).

Supplementary Fig. 20. Inhibition of caseinolytic activity and substrate specificity of Atlantic and Baltic hatching enzymes. (A) Inhibition of caseinolytic activity of Atlantic and Baltic hatching enzymes. Activity is expressed as relative activity, with the activity in the absence of inhibitor (–) defined as 100%. The symbol (+) indicates the presence of EDTA. The partially purified enzyme was incubated with 10 mM EDTA solution for 10 min at 30 °C before testing caseinolytic activity as described above. (B, C) The substrate specificity of the partially purified hatching enzymes of Atlantic (B) and Baltic (C) herring examined using

eighteen 4-methylcoumaryl-7-amide (MCA) peptides. The sequences of MCA peptides are shown on the X-axis. Activity toward each substrate is expressed as a percentage relative to the substrate showing the highest activity (set as 100%). MCA peptide cleavage activity was measured using a 50 μ L reaction mixture consisting of 50 mM Tris-HCl (pH 8.0), 100 μ M MCA peptide (Peptide Institute, Osaka, Japan), NaCl (0.25 M for Atlantic herring and 0 M for Baltic herring), and the enzyme. The mixture was incubated for 1 h at 30°C. The reaction was terminated by adding 100 μ L of 20% acetic acid. Fluorescence intensity of the solution was measured using a plate reader with excitation at 360 nm and emission at 460 nm. AMC standards (serial dilutions from 10 mM AMC, 1:1 to 1:1024) were used for calibration.

Supplementary Table 1. Regions of genetic differentiation between Atlantic and Baltic spring-spawning herring. For each region, the section, if any, overlapping with the Ringkøbing vs Atlantic contrast is indicated, as are the genes highlighted in Fig. 1.

Supplementary Table 2. Nomenclature for *LRRC8C*, *ZPB1*, *FTG* and *HE* genes in the Atlantic herring genome.

Supplementary Table 3. Species and genomic resources used for *LRRC8C* analyses.

Supplementary Table 4. Overview of LRRC8C1 and LRRC8C2 peptides monitored in sperm and tissue samples, whether presence could be verified and the corresponding limit of detection determination.

Supplementary Table 5. The selected proteins from the Atlantic herring sperm depicted in the ranking plot (Fig. 2e). iBAQ values and other information are taken from the MaxQuant output table (Dataset S1).

Supplementary Table 6. Pore residues in mouse LRRC8C and herring LRRC8C1 and LRRC8C2 corresponding to residues K98, D100, R103, and H104 in human LRRC8A.

Supplementary Table 7. Sequences from other vertebrate species homologous to *FTG* and *F13A* genes in Atlantic herring, used in the phylogenetic analysis.

Supplementary Table 8. Average d_N , d_S and d_N/d_S ratios for the *FTG* and *F13A* genes in different branch groups based on Fig. 3a.

Supplementary Table 9. The selected proteins from the Atlantic herring oocyte depicted in ranking plot (Fig. 3e). iBAQ values and other information are taken from the MaxQuant output table (Dataset S1)

Supplementary Table 10. Results of *in vitro* fertilization of Atlantic and Baltic herring at different salinities.

Supplementary Table 11. Protein density of egg envelope determined by egg diameter, thickness of inner layer and protein amount per egg.

Supplementary Table 12. Hatching enzyme sequences from teleost species used as queries for gene annotation and for phylogenetic reconstruction.

Supplementary Table 13. Analysis of copy number variation at the *HEIC* locus on chromosome 26 in different herring populations based on short read, pooled whole genome resequencing.

Supplementary Table 14. Chromatographic separation used for PRM LC-MS/MS analysis of sperm and tissue samples.

Dataset 1. Mass spectrometry data of sperm (a), oocytes, (b), and the tissues brain, heart, lung, skeletal muscles, and spleen (c). For sperm and oocytes MaxQuant output tables and for the tissues Spectronaut output tables are given.

Materials and Methods

Nomenclature

None of the genes characterized in the present study is properly annotated in the current Atlantic herring reference genome assembly Ch_v2.0.2v2 (GCA_900700415.2). We have therefore compiled Table S2 that provides the current annotation and our proposed revised nomenclature.

Genetic screen based on whole genome sequencing data

The per-population frequencies used for the delta allele frequency contrasts (Figures and heatmaps stem from previously published data sets from Han *et al.* (1) and Goodall *et al.* (2) (<https://doi.org/10.17044/scilifelab.22361761.v1> and EVA:PRJEB82679). Delta allele frequency was calculated as the absolute difference of the means, with each constituent pools given equal weight, of the two contrasted groups. The averages used comprised the following pools (names according to Supplementary File 1 of Han *et al.* (1)).

Atlantic ($n = 18$): DalFB_Atlantic_Spring, DalInB_Atlantic_Spring, DalNsS_Atlantic_Spring, HGS15_NSSH_Atlantic_Spring, HGS20_CapeWrath_Atlantic_Spring, HGS23_Clyde_Atlantic_Spring, HGS24_Landvik_Atlantic_Spring, HGS25_Lindas_Atlantic_Spring, HGS26_Lusterfjorden_Atlantic_Spring, HGS27_Gloppen_Atlantic_Spring, HGS8_KattegatNorth_Atlantic_Spring, HGS9_Greenland_Atlantic_Spring, LandvikS17_Norway_Baltic_Spring, O_Hamburgsund_Atlantic_Spring, PB10_Skagerrak_Atlantic_Spring, PB2_Iceland_Atlantic_Spring, PB9_Kattegat_Atlantic_Spring, Q_Norway_Atlantic_Atlantic_Spring.

Baltic ($n = 15$): A_Kalix_Baltic_Spring, B_Vaxholm_Baltic_Spring, G_Gamleby_Baltic_Spring, HGS1_Riga_Baltic_Spring, HGS2_Riga_Baltic_Spring, HGS6_Schlei_Baltic_Spring, HGS71_Rugen_Baltic_Spring, HGS72_Rugen_Baltic_Spring, J_Traslovslage_Baltic_Spring, PB11_Kalmar_Baltic_Spring, PB12_Karlskrona_Baltic_Spring, PB1_HastKar_Baltic_Spring, PB4_Hudiksvall_Baltic_Spring, PB5_Galve_Baltic_Spring, PN3_CentralBaltic_Baltic_Spring

Ringkøbing ($n = 1$): HGS11_RingkøbingFjord_NorthSea_Spring

For definition of independent regions of differentiation, the criterion was to retain SNPs with delta allele frequency > 0.5 in the relevant contrasts, followed by joining of all SNPs with their nearest neighbour at most 100 kb away. This procedure captures the top of all qualifying signals, but typically the differentiation is notably elevated over a larger footprint. See Han *et al.* (1) for a more exhaustive list of differentiated markers between Atlantic and Baltic herring.

For all heatmaps, the SNPs to be included were selected based on a delta allele frequency cut-off determined based on the height of the peak at the locus in question. In order, these cut-offs were: 0.5 (*LRRC8C2*), 0.75 (*FTG*), 0.5 (*ZPBA1*), 0.5 (*HEI*).

Genome-wide case-control association analysis was performed to contrast Baltic Sea and Atlantic Ocean herring populations using PLINK (v1.9) (3). Individuals were assigned to target populations and recoded into binary case-control status based on sampling origin. Variants were filtered to exclude SNPs with $>10\%$ missing genotypes (`--geno 0.1`) and minor allele frequency <0.1 (`--maf 0.1`). Association testing was conducted using PLINK's allelic case-control test (`--assoc`), with population identity coded as control (1) and case (2). Genomic inflation was assessed and corrected using genomic control (`--adjust gc`). Genome-wide significance was determined using a Bonferroni correction based on the total number of tested variants, and results were visualized using Manhattan and QQ plots.

Individual haplotype analysis

As was the case for the population frequencies above, the analysis of individual haplotypes was performed on previously published data (1, 4). The distances were calculated in R using the “stringdist” package v0.9 (<https://github.com/markvanderloo/stringdist>), while the trees were made using the “bionj” function from the “ape” package v5.3 (5).

Proteomic analysis of sperm and egg

Sample preparation for LC-MS/MS analysis of sperm and oocytes. Baltic herring was collected from commercial fishing ~80 km north of Uppsala, Sweden (60° 38' 52.0'' N, 17°48' 44.2'' E), on June 19, 2023. Egg and sperm were prepared and transported in cooling boxes to the Max-Planck-Institute for Multidisciplinary Sciences in Göttingen., Germany.

Sperm samples (n=3) were lysed as described (6) and yielded between 1.7 mg and 3.6 mg protein per sample. Oocytes (n=3) were lysed with 4% (w/v) SDS, 150 mM NaCl, 150 mM Hepes/NaOH pH 7.5, 2 mM DTT, 0.5% (w/v) NP40, 4 mM EDTA, and 1x Roche complete protease inhibitor/EDTA; the yield was between 7 mg and 8.5 mg protein per sample. Further sample processing of sperm and oocyte samples was performed as described (6), including sonification (Bioruptor device Plus, Diagenode), incubation with universal nuclease (Pierce 250 U/μl), reduction and alkylation of sperm proteins. Between 450 and 250 μg of sperm proteins per sample and 500 μg to 1 mg of oocyte proteins per sample were aggregated on SP3 beads (Cytiva) for sample clean up (7). Digestion was performed in 50 mM NH₄CO₃ with 0.1 % (w/v) Rapigest (Waters) and endoproteinase trypsin 1:20 (w/w) overnight at 37° C as described (6). All further steps of peptide extraction and C18 spin-column desalting (Harvard Apparatus, Holliston, MA) of peptides were also performed as previously described (6). Peptides (100 μg) derived from each of the three sperm and oocytes protein samples were separated by off-line basic RP-HPLC (bRP-HPLC) using an Agilent 1100 series as described (6). Twelve combined bRP-HPLC fractions with peptides from each sperm and oocyte sample were dried in a SpeedVac concentrator for LC-MS/MS analysis.

Sample preparation for LC-MS/MS analysis of tissue samples. Roughly 1 mm³ tissue samples corresponding to 1-2 mg wet weight were lysed in 30 μl freshly prepared lysis buffer (Complete protease inhibitor cocktail (Roche), PhosSTOP phosphatase inhibitor cocktail (Roche), 20 mM TCEP, 50 mM IAA, 2% SDS in aqueous 100 mM HEPES pH 8.0) under pressure cycling conditions in a Barocycler 2320XT (Pressure BioSciences). 60 pressure cycles were used consisting of 50 s at 45 kpsi followed by 10 s at normal pressure, respectively. Following centrifugation for 2 min at 5000x G to remove debris, protein concentrations were determined by a microBCA assay on a Nanodrop 2000 UV/Vis Spectrophotometer (Thermo Fisher Scientific). Aliquots volume-adjusted to 15 μg protein content were then reconstituted to 50 μl volume with lysis buffer, absorbed onto paramagnetic beads (Resyn Amine, Resyn Biosciences), washed and tryptically digested using a modified SP3 protocol (8) on a Kingfisher Duo Prime magnetic bead handler (Thermo Fisher Scientific). Samples were acidified using a 1% solution of trifluoroacetic acid in 20% aqueous acetonitrile and containing iRT standard peptides (Biognosys, Schlieren, Switzerland) to a final concentration of 0.1% TFA and 2% acetonitrile, respectively. Peptide concentrations were measured using a nanofluidic UV/Vis spectrophotometer (Little Lunatic, Unchained Labs), and the samples then directly used for mass spectrometric analysis without further processing.

LC-MS/MS analysis of sperm and oocytes. Dried peptides of the bRP-HPLC fractions were dissolved in 4% (v/v) acetonitrile/0.1% trifluoroacetic acid in H₂O. Twelve fractions of each sperm and oocyte sample were analysed by LC-MS/MS on an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific) coupled to a Dionex UltiMate 3000 UHPLC system

(Thermo Fisher Scientific). Peptides were separated on an in-house packed analytical column (30 cm length; ReproSil-Pur 120 Å, 3 µm pore size, C18-AQ; 75 µm inner diameter) with a gradient of 5-8% solvent B (0.08% (v/v) formic acid/80% (v/v) acetonitrile) in 7 min, 8-35% B in 90 min, 35-55% B in 7 min and 90% B in 4 min at a flow rate of 300 nL/min. MS and MS/MS data were acquired in a data-dependent acquisition (DDA) mode using a 3-s cycle time. Full MS scans were recorded at a resolution of 120,000 (at m/z 200) over an m/z range of 350–1600. The AGC target was set to 300% with a maximum injection time of 25 ms. The RF lens was operated at 50%, and source fragmentation was disabled. MS/MS spectra were acquired for selected precursors using higher-energy collisional dissociation (HCD). Precursors were isolated with a 1.4 m/z isolation window. Fragment ions were analysed in the Orbitrap at a resolution of 15,000 FWHM in centroid mode. The normalized higher-energy collisional dissociation (HCD) was 28%. The AGC target for MS/MS scans was set to 75% with a maximum injection time of 80 ms.

LC-MS/MS analysis of tissue samples. Mass spectrometric analysis with data-independent acquisition was performed using a nanoflow chromatography system (nanoRSLC, Thermo Fisher) hyphenated to a hybrid timed ion mobility-quadrupole-time of flight mass spectrometer (timsTOF Pro 2, Bruker). Briefly, 400 ng of peptide equivalents were enriched on a reversed-phase C18 trapping column (0.3 cm × 300 µm PepMap C18, Thermo Fisher) and separated on a reversed-phase C18 column equipped with an integrated CaptiveSpray Emitter (Aurora 25 cm × 75 µm, IonOpticks). Peptide separation was performed with a 100-minute linear gradient of 5-34% acetonitrile containing 0.1% formic acid (v/v) at a flow rate of 200 nL/min and a column temperature of 50°C.

Data-independent acquisition analysis was performed in diaPASEF mode (9) using a 20x2 variable size window acquisition method from m/z 400 to 1,200 to include the 2+/3+/4+ population in the m/z -ion mobility plane. The collision energy was incrementally and linearly increased as a function of mobility from 59 eV at $1/K_0=1.5\text{Vs cm}^{-2}$ to 20 eV at $1/K_0=0.7\text{Vs cm}^{-2}$. Two mass spectrometry replicates per biological replicate were acquired.

Sample preparation for PRM LC-MS/MS. Heavy isotope-labelled standard peptides (followingly denoted as heavy peptides) from ThermoFisher (AQUA QuantPro) were received solubilized with a concentration of 5 pmol/µL in 5% (v/v) ACN in H₂O. Heavy peptides for LRRC8C1 and LRRC8C2 were mixed at equimolar concentrations and split into aliquots of 10 pmol, dried in a SpeedVac concentrator and stored at –20 °C. For each targeted analysis, a 10 pmol aliquot of heavy peptides was used. Heavy peptides were resuspended in 4% (v/v) ACN/0.1% (v/v) TFA in H₂O and spiked into sperm and tissue samples at different amounts to finally reach 0.005, 0.01, 0.05, 0.1, 0.5, 1.0, 5.0, 10.0, and 50.0 fmol of heavy peptide per 7 µL, which was the injection volume for LC-MS/MS analysis. Sperm and tissue samples digested as described before were resuspended directly in the heavy peptide solutions yielding 1 µg of peptides in 7 µL. Samples were analyzed in biological triplicates.

PRM LC-MS/MS analysis of sperm and tissue samples. Mass spectrometric analysis with parallel reaction monitoring was performed on an Orbitrap Exploris™480 mass spectrometer (Thermo Fisher Scientific) coupled to a Dionex UltiMate 3000 UHPLC system (Thermo Fisher Scientific) equipped with an in-house-packed C18 column (ReproSil-Pur 120 C18-AQ, 1.9 µm pore size, 75 µm inner diameter, 30 cm length, Dr. Maisch HPLC, Ammerbuch, Germany). The LC-MS system was operated using the software Thermo XCalibur 4.4.16.14 for the Orbitrap Exploris™480 mass spectrometer. Peptide separation was performed using the gradient given in Table S14; 0.1% (v/v) FA in H₂O was used as mobile phase A and 80% (v/v) ACN/0.08% FA (v/v) in H₂O as mobile phase B.

During PRM measurement, MS1 and targeted MS2 scans were performed, both in positive ion mode. For MS1 acquisition, spectra were recorded in the Orbitrap at a resolution of 120,000 (FWHM). The m/z scan range was adjusted to encompass the expected mass range of the target peptides. A standard AGC target was applied. For targeted MS2 analysis, precursor ions were isolated with a 1.4 m/z window and fragmented using a normalized collision energy of 30%. The normalized AGC target was set to 100%, with a maximum injection time of 100 ms. MS2 spectra were acquired with a resolution of 60,000 (FWHM). Target selection was based on predefined precursor m/z values, charge states, and retention time windows specified in the PRM method. Retention time windows were determined individually for each measurement series. For each sample, 1 µg of peptide were measured per run.

Data analysis of DIA LC-MS/MS

Database search and data analysis. Protein identification and quantitation were achieved in Spectronaut Software v20.2 (Biognosys) (10) in directDIA mode. For identification, the in-built Pulsar algorithm was used at default settings against a UniProtKB *Clupea harengus* reference proteome (revision 08-2023) augmented with the sequences of eight LRCC8C protein alleles and an in-house database of 54 frequently identified laboratory contaminants. Precursor, peptide and protein identifications were all adjusted to 1% FDR using a forward-and-reverse decoy database approach. Quantification involved selecting up to 6 of the most abundant fragment ion traces for each peptide and up to 10 peptides per protein, integrating and summing them to calculate protein area values. Quantitation values were globally normalized and missing values imputed cross-run using default settings.

Data analysis of DDA LC-MS/MS

Database Search. The MS raw files of the LC-MS analysis of sperm and oocytes were imported into MaxQuant (version 2.7.5.0) (11, 12) and searched with Andromeda search engine (13) against the UniProt *Clupea harengus* database with 41,031 entries (released on November 24, 2025) plus three FASTA sequences of LRRC8C2, Baltic North, Baltic South, and Intermediate. Trypsin/P was set as digestion enzyme with maximal two missed cleavage sites. A minimum peptide length of 7 amino acids was required for protein identification. Mass tolerances were set to 20 ppm for the first search, 4.5 ppm for the main search, and 20 ppm for FTMS MS/MS fragment ions. Carbamidomethylation of cysteine was set as a fixed modification, and oxidation (M), acetylation (protein N-terminus), were set as variable modifications. MaxQuant built-in contaminants list was included and peptide-spectrum match, peptide, and protein FDR thresholds were set to 1%. For quantification of sperm and oocyte proteins, the iBAQ option (14) in MaxQuant was enabled.

Downstream Data Analysis. Proteins in MaxQuant output tables were filtered to remove entries annotated as reverse hits and contaminants. Data visualization, i.e., ranking plot analysis based on iBAQ values from the MaxQuant output was performed with ggplot2 (version 4.0.0) and eulerr (version 7.0.4) on R studio (version 4.3.1). Sequence alignment was done with ClustalX tool available in Uniprot.com with default settings.

Data analysis of PRM LC-MS/MS

Data Analysis. The PRM data was analyzed with Skyline (MacLean et al., 2010, 25.1.0.237, MacCoss Lab Software, University of Washington, Seattle, WA). The mass accuracy (delta mass ≤ 10 ppm) of the transitions was used as cut-off filter. The program was provided with

peptide sequences; in the peptide settings section, the isotope modifications for ^{13}C , ^{15}N -labelled lysine and arginine were added. Acquisition method was specified as PRM, product mass analyzer as Orbitrap and the resolving power as 60K. The MS raw files were imported into Skyline. Transitions of endogenous peptides were compared to signals of labelled peptides; presence of endogenous peptides was considered verified when they showed the same retention time as the labelled peptides and similar fragmentation pattern meaning measurement of at least three out of four selected fragment ions and sharing the same fragment ion as the most abundant.

Analysis of PacBio long-read data

PacBio data were available for 15 individuals, eight Atlantic herring and seven Baltic herring (15, 16). Six of the Atlantic herring were collected from the Celtic Sea (latitude 51°59'N, longitude 6°48'W; salinity 35 ppt) and two from the Norwegian Sea (latitude 67°55'N, longitude 11°18'E; salinity 35 ppt). The Baltic herring included seven individuals from Hästkär (latitude 60°35'N, longitude 17°48'E; salinity 6 ppt). One of these samples has been used to construct a new Atlantic herring reference genome (Ch_v3.1, GCA_040183275.2). The data was generated using PacBio HiFi technology with sequencing depth ranging from 23-30x (15). *De novo* genome assemblies were generated using Hifiasm (v0.16.1-r375) (17), which produced 30 haplotype-resolved assemblies by separating each genome into primary (hap1) and secondary (hap2) haplotypes. Assembly procedures and summary statistics have been described (15). Contig-level assemblies were scaffolded into chromosome-level assemblies using RagTag (v2.0.1) (18) with the Atlantic herring reference genome (Ch_v2.0.2v2). The resulting chromosome-level, phased assemblies were used in downstream analyses.

Phylogenetic analysis of molecular evolution

Leucine-rich repeat-containing family 8 isoform C (LRRC8C)

Assessments of *LRRC8C* evolution spanned clupeid genome data sampled across three families (Table S3). Only the genomes of *D. clupeioides*, *C. harengus*, *A. alosa* and *A. sapidissima* had publicly available annotations of protein coding genes at the time of analysis. We therefore used spaln (v2.4.6) (19) with parameters “-Q7 -LS -O0,1,2,3,4,6,7,15” to perform spliced alignments using non-redundant sets of protein sequences of two annotated clupeids to infer gene models across the unannotated genomes. To this end, we used *C. harengus* (n=24,095 proteins) to map genes in *C. nasus* and *S. sprattus* and *A. sapidissima* (n=25,717 proteins) to map genes in *S. pilchardus*, *L. miodon* and *T. ilisha*. For each newly annotated species, duplicate gene models with coding DNA sequence identity higher than 99.5% were removed using CD-HIT with parameters “cd-hit-est -c 0.995 -n 8 -M 0 -aS 0.80 -G 0 -g 1”, to preemptively control for possible redundancy due to cryptic haplotigs. We used BUSCO (v5.7.1) (20) to assess the overall completeness of the gene annotations, which spanned 21–24 K genes and 78–99% BUSCO completeness per species and low duplication levels (Table S3).

To detect *LRRC8C* genes across species, we first carried out proteome-wide orthogroup clustering with ProteinOrtho (v6.3.1) (21) using DIAMOND (v2.0.4) (22) to perform queries and OrthoSNAP (23) to split large clusters. We detected two separate orthogroups for *LRRC8C1* and *LRRC8C2*, respectively. We then manually performed DIAMOND searches among species for missing sequences and detected an annotated model for *LRRC8C1* in *D. clupeioides* (ENSDCDG00000028838.1) that had been missed by the automated clustering, but could not find evidence of *LRRC8C2* in either *D. clupeioides* or *C. nasus*. Because *LRRC8C* sequences were nearly identical between *A. alosa* and *A.*

sapidissima, we excluded the former from downstream phylogenetic analyses, as limited numbers of substitutions between sequences make it difficult to infer signatures of selection. We thus used eight *LRRC8C1* clupeid sequences and six *LRRC8C2* clupeid sequences to study molecular evolution. In the case of *C. harengus*, ‘Atlantic’ alleles of *LRRC8C1* and *LRRC8C2* were used. To this dataset, we added *LRRC8C* orthologs of other fish and mammals. Most of these sequences were reported as orthologs of the Atlantic herring *LRRC8C1* (ENSCHAG00020014387) in Ensembl and downloaded as a single set (n=119). The set included two sequences from the Atlantic herring and one from the Denticle herring (*D. clupeoides*) but no other clupeid sequences. These were replaced by the *LRRC8C* sequences detected above. To the dataset, we added NCBI *LRRC8C* sequences for haddock (OZ180136.1) and burbot (OZ177867.1) with the aim to shorten a long internal branch to clupeids, while pruning the Ensembl set to include 13 representative and well-aligning fish sequences and four outgroup mammalian sequences, in order to make the overall dataset practical for PAML analyses. The completed dataset contained 33 *LRRC8C* sequences, including duplicate sequences from goldfish and common carp, which share a recent lineage-specific whole-genome duplication (24).

We used the local version of TranslatorX (25) and MAFFT (v7.407) (26) to align the *LRRC8C* coding sequences at the protein level to ensure that open reading frames were retained. We then inferred a Bayesian gene tree using MrBayes (v3.2.7a) under the GTR+ Γ model (“lset nst=6 rates=gamma;”) using uninformative priors for substitution parameters and base frequencies, with two independent 10-chain MCMC runs for 2 million generations each, sampling every 100 generations. The first 2,500 samples (i.e. 250 thousand generations) were discarded as burn-in, after which the average standard deviation of split frequencies was <0.005 throughout the rest of the analysis, indicating stationary posteriors and convergence among chains. We produced a consensus tree and re-rooted it on the node separating fish and mammals using Phyx (v1.3) (27).

Lastly, we performed a phylogenetic analysis of molecular evolution of *LRRC8C* using codeml in PAML (v4.9j) (28) through ETE Toolkit (29) with the aim to infer signatures of selection among the clupeid sequences. We used a free-ratio branch model to estimate nonsynonymous/synonymous rate ratios ($\omega = dN/dS$) for each branch across the tree, assuming no site heterogeneity, no molecular clock and codon frequencies estimated under the F3x4 model, while estimating both κ (transition/transversion ratio) and ω under maximum likelihood. We extracted the tree with branch-lengths scaled by ω from the PAML output and plotted it with ggtree (release 3.21) (30) in R (v4.5).

Fish transglutaminase (FTG)

Fish transglutaminase (FTG) gene originated from duplication of *factor 13 subunit A (F13A)* gene in the ancestor of Teleostei (31). However, in current genome annotation databases, this gene is frequently annotated as *F13A* or remains unnamed. In order to identify *F13A* and *FTG* genes in the Atlantic herring genome, we performed a homology-based search using orthologous sequences from other teleost species. We identified two candidate genes on chromosome 15, three on chromosome 17, and one on chromosome 18. By genomic synteny comparisons with other teleosts, we found that the copy on chromosome 18 shares conserved synteny with *F13A* in other species, while the copy on chromosome 17 aligns with *FTG*. The genes on chromosome 15 did not exhibit conserved synteny. Subsequent phylogenetic analysis, described in the next paragraph, confirmed that these chromosome 15 copies clustered closely with the *FTG* lineage. Ensembl (Release 114) annotations were used as initial gene models for *F13A* and *FTG* and were manually curated using validated orthologous sequences from rainbow trout, medaka, and zebrafish to improve model accuracy.

Phylogenetic analysis was performed for *FTG* and *F13A* genes from Atlantic herring and other fish, with *F13A* sequences from three mammalian species included as outgroups. Sequences from other species were retrieved from public databases, with accession numbers provided in Table S7. For species lacking annotation in public databases (*Limnothrissa miodon*, *Sprattus sprattus* and *Tenualosa ilisha*), gene models generated in this study (see above) were used (Table S3). For Atlantic herring, representative Atlantic alleles of *FTG* were included in the analysis. Amino acid sequences were aligned using MAFFT (v7.407) (26), and the resulting alignment was used to generate codon-based nucleotide alignments with PAL2NAL (v14) (32), preserving reading frames and codon structure. Maximum-likelihood phylogenetic tree was inferred using IQ-TREE (v2.2.2.6) (33, 34) under the codon substitution model GY+F+I+R4, with branch support assessed using 1,000 ultrafast bootstrap replicates (35). The resulting tree was re-rooted using R package ape (v5.8-1) on the branch leading to the mammalian *F13A* clade.

To investigate patterns of molecular evolution and selection of *FTG* and *F13A* genes, we employed codeml from PAML (v4.10.7) (28) using the free-ratio model (model = 1), which allows each branch in the tree to have its own d_N/d_S ratio (ω). We used the maximum-likelihood tree topology inferred using IQ-TREE. The codon frequencies were estimated using the F3×4 model, with analysis conducted under assumptions of no site heterogeneity (NSsites = 0) and no molecular clock (clock = 0). The analysis yielded d_N , d_S , and d_N/d_S estimates for each branch. Branch lengths of the maximum-likelihood tree were rescaled using the estimated d_N values, and the resulting d_N -scaled tree was visualized in R using (v4.4.3) ggtree (v.3.12) (30).

After establishing gene models on the reference genome (Ch_v2.0.2v2), *FTG* genes were annotated in all 30 PacBio assemblies using LiftOff (v1.6.3) (36), followed by manual curation to ensure intact open reading frames and accurate splice-site annotation. All assemblies contained five *FTG* gene copies, of which disrupted (putative null) copies were excluded from downstream analyses. To further resolve the evolutionary relationships among *FTG* copies within Atlantic herring, we conducted a separate phylogenetic analysis focusing on these gene copies. Coding sequences (CDS) of all five *FTG* genes were extracted from each PacBio assembly using GffRead (v0.12.7) (37) and combined with *FTG* sequences from the reference genome. A single *FTG* sequence from European sprat was included as an outgroup. Sequences were aligned using MAFFT, and a maximum-likelihood phylogenetic tree was inferred using IQ-TREE with the MG+F3X4+R2 codon substitution model. Tree visualization was performed using the iTOL online platform (38). Nucleotide diversity (π) was calculated separately for Atlantic and Baltic groups for the three *FTG* gene copies on chromosome 17 using the R package pegas (v1.3)(39).

Hatching enzyme (HE)

The reference genome assemblies for Atlantic herring and European sprat were retrieved from the NCBI Assembly database. For Atlantic herring, we used Ch_v3.1 (GCA_040183275.2) because this most recent assembly is based on a Baltic herring. For European sprat, we used the fSprSpr1.1 principal haplotype (GCA_963457725.1; Whitsand Bay, Cornwall, UK). Hatching enzyme (HE) sequences previously reported from other teleost species were used as queries in the gene annotation process (40). Protein query sequences were aligned to the reference genomes using Spaln (19) for splice-aware mapping. Alignment hits were inspected in IGV (v2.19.7) (41), and the best hit per gene locus was selected. Selected gene models were manually curated to ensure the completeness of each gene in terms of exon boundaries, canonical splice sites, and the integrity of the open reading frame to accurately define gene structure. To facilitate identification and referencing of tandem gene copies, we designated

each copy by its chromosome and genomic midpoint position (in Mb), formatted as species_chr#_position (e.g., *Clupea harengus*_chr26_6.277).

We used conserved synteny and phylogenetic analyses to distinguish *HE* genes from other C6astacin paralogues, in order to exclude other C6astacin family members from downstream analyses. C6astacin genes other than *HE* typically exhibit relatively conserved synteny across species within clades, in contrast to the highly dynamic genomic organization of *HE* genes. Synteny analysis was performed by identifying candidate genes based on their proximity to genes within previously defined conserved synteny sets, as described (42).

For phylogenetic analysis, coding sequences (CDS) of *HE* genes from Atlantic herring and European sprat were extracted from their respective reference genomes using GffRead (37). Representative *HE* orthologues from other teleosts were obtained based on accession numbers listed in Kawaguchi *et al.* (40), and additional C6astacin sequences were similarly obtained from Nagasawa *et al.* (42). The sequences were then retrieved from NCBI using these accession numbers. Only the coding region corresponding to the mature enzyme was retained and the flanking sequences were trimmed. Amino acid sequences were aligned using MAFFT (26), and codon-based nucleotide alignments were generated with PAL2NAL (32). A maximum-likelihood phylogenetic tree was inferred using the MG+F3X4+R5 codon substitution model with IQ-TREE (33, 34), with 2,000 ultrafast bootstrap replicates (35) to assess branch support. Tree visualization was performed using ggtree package (v3.12)(30) in R (v4.4.3).

Based on phylogenetic and synteny analysis, C6astacin paralogues other than *HE* were identified and excluded from further analysis as they were only used for preliminary exploratory analyses. For the final phylogenetic tree, only verified *HE* sequences were retained. A list of *HE* sequences from other teleost species used in the annotation and phylogenetic reconstruction, sourced from Kawaguchi *et al.* (40), is provided in Table S12.

Amino-acid sequences of *HE1* genes were extracted from the Atlantic herring reference genome using GffRead, and aligned with MAFFT. Alignments were trimmed to include only the mature protein region. For the alignment plot, polymorphic sites were filtered and visualized using custom R script.

Annotation and copy number analysis of *HE1C* genes in PacBio assemblies

We annotated the *HE1C* gene copies on chromosome 26 across all PacBio assemblies to investigate their organization and copy number variation in the Atlantic and Baltic haplotypes. For initial annotation, Spaln (v3.0.3) (19) was used to align protein query sequences from the corresponding region in the Atlantic herring reference genome (Ch_v3.1) to each haplotype assembly. To ensure complete detection of tandemly duplicated *HE1C* gene copies, including those potentially missed during initial annotation, we applied a complementary approach. Specifically, Spaln alignments were repeated using the hatching enzyme protein query set originally used for annotation of the reference genome. Additional copies identified through this step were manually added to the annotation set by selecting the best alignment hit per gene locus. All predicted gene models were examined and manually curated in IGV (41), following the same procedure used for the reference genome, to ensure the integrity of exon-intron boundaries, open reading frames, and correct gene structure.

In addition, we annotated neighbouring genes at the *HE1C* locus, including *HSP70*, *DNAJA3B* and *CCNBI*, in all PacBio assemblies. Coding sequences and corresponding amino acid sequences were retrieved from the Atlantic herring reference genome annotation (Ch_v2.0.2v2; GCA_900700415.2) available in Ensembl (Release 114) and used as queries in Spaln alignments. To recover additional *HSP70* copies that may have been missed by splice-aware alignment, we also applied MUMmer (v4.0.0rc1) (43) and BLAST+ (v2.14.1) (44, 45), aligning the query sequences to each PacBio haplotype. The outputs from NUCmer and

BLAST were converted to BED format and visualized alongside Spaln results in IGV for manual curation.

Gene annotations were merged into a concatenated GFF3 file per assembly and visualized using gggenomes (v1.0.1) (46) in R (v4.4.3). In three haplotypes (BS4_hap1, BS4_hap2 and BS6_hap2), the *HEIC* locus was split across two contigs, introducing a break in haplotype continuity. In all other cases, the locus was fully spanned by a single contig without such disruption.

Correlation between *HEIC* gene copy number and salinity across populations

Relative copy number of *HEIC* genes on chromosome 26 was estimated using previously reported pooled sequencing data (see above) using a read-depth-based approach. Sequencing depth was calculated with SAMtools depth (v1.20) (47) across the whole genome and across the genomic interval spanning the three tandem *HEIC* gene copies in the reference genome (Ch_v2.0.2v2). For each pooled sample, relative copy number was inferred as three times the ratio of the average read depth across the *HEIC* locus to the average genome-wide read depth. The *HEIC* copy number estimates were then tested for correlation with local salinity using linear regression in R (v4.4.2). Salinity data were obtained from the Copernicus Marine Service database

(<https://data.marine.copernicus.eu/products?facets=specificVariables%7ESalinity>) based on the recorded longitude, latitude, and sampling date for each pool. To minimize potential spatial and temporal biases, the salinity for each sampling site was calculated as the average value within $\pm 0.1^\circ$ in both longitude and latitude, and averaged across measurements taken within ± 15 days of the sampling date. This approach yields a robust local salinity estimate reflecting both spatial and temporal environmental conditions.

***In vitro* fertilisation and collection of materials from egg and embryos**

Prior to carrying out these experiments, the local animal welfare committee at the Department of Biosciences, University of Bergen, confirmed that the experiments on unfed herring embryos were within national guidelines and legislation and exempted from further permit application. Atlantic herring samples for *in vitro* fertilization were obtained ~12 km west of Bergen, Norway (60°34' 11.2''N, 5° 0' 18.9'' E) on the morning of April 7, 2025, and Baltic herring, in late evening ~80 km north of Uppsala, Sweden (60° 38' 52.0'' N, 17°48' 44.2'' E), on May 5, 2025. Herring were transported in cooling boxes to the laboratory at University of Bergen, within 2 hours (Atlantic) and 12 hours (Baltic) after retrieval from the nets. Fish measurements and crossings were completed within 2 hours upon arrival. Mixed sperm samples from three males per population (Atlantic and Baltic) were activated in 16‰ salinity in both cases and used to fertilize eggs of single females on glass plates, microscope slides or NUNC trays at respective salinities (see Berg et al. (48) for details on procedures). Mean male size averaged 34.3 cm and 329 g, and 20.3 cm and 62 g for Atlantic and Baltic males, respectively. Size of females used in crossings averaged 30.3 cm and 248 g, and 19.5 cm and 57 g for Atlantic and Baltic females, respectively. The eggs were incubated at nominal salinities of 6, 16 and 35‰, and monitored daily until hatching. Crosses that were unsuccessful at all tested salinities, were excluded from further analysis as this was most likely a result of long postmortem duration from sampling to stripping. In total crossings with four Atlantic and eleven Baltic females were available for further analyses, with fertilization rates ranging from 2 to 95% (Table S10). Samples of developing eggs and newly hatched larvae were flash frozen in liquid nitrogen, fixed in PFA/PBS or ethanol fixatives. Frozen samples were stored in -80 °C freezers and shipped on dry ice or at -18°C to laboratories in Sweden and Japan for subsequent analyses.

RNA-seq analysis

A total of 18 samples was used for RNA-seq analysis, three replicates of Atlantic and Baltic herring, at three different developmental stages. Total RNA was extracted with the Direct-zol™ RNA MiniPrep kit (Zymo Research). RNA quantity and quality were assessed with a TapeStation 4150, RNA ScreenTape and RNA ScreenTape Sample Buffer (Agilent Biotechnologies). The quality of the RNA was determined by the RNA Integrity Number (RIN); a RIN ≥ 8 is considered as the minimum standard value. Twenty-one out of twenty-eight samples gave the desired RIN. RNA samples were processed by the SNP&SEQ Technology Platform at Uppsala university for library preparation and sequencing. There, RNA integrity was probed again, this time with a Fragment Analyzer in combination with the DNF-471 Standard Sensitivity RNA kit (Agilent Biotechnologies); the quality output is in this case the RNA Quality Number (RQN). Six samples were also double-checked at a TapeStation prior to library construction. Two-hundred ng of total RNA were used as input for the TruSeq stranded mRNA library preparation kit with unique dual indexes (Illumina). The libraries were sequenced on a NovaSeq X Plus system with a 10B flowcell and XLEAP-SBS. The chemistry was paired-end 150 bp read length.

Raw sequencing reads were inspected for overall quality and adaptor contamination using MultiQC (v1.22.2) (49) and Trim Galore (v0.6.7). Residual adaptor sequences and low-quality bases were trimmed, and reads shorter than 30 bp in length were discarded. Clean reads were aligned to the Atlantic herring reference genome (Ch_v2.0.2v2) using HISAT2 (v2.2.1) (50) with default settings. Alignment files in SAM format were converted to BAM, sorted and indexed using SAMtools (v1.20) (51).

Aligned reads were quantified at the gene level using StringTie (v2.2.1) (52) in estimation mode guided by the reference annotation (Ensembl release 115, https://ftp.ensembl.org/pub/release-115/gtf/clupea_harengus/). For each sample, StringTie produced gene-level expression estimates, and gene-level count matrices were generated using the *prepDE.py* script distributed with StringTie (52). The resulting raw integer counts per gene were used as input for all downstream differential expression analyses.

Differential gene expression between Atlantic and Baltic herring was assessed using DESeq2 (v1.46.0) (53) in R (v4.4.2). The raw gene count matrix was imported into DESeq2 (53), and size-factor normalization and dispersion estimation were performed using default settings. Negative binomial generalized linear models were fitted for each gene, and differential expression for each comparison (e.g., Baltic vs Atlantic at different developmental stages) was evaluated using Wald tests (53). Genes with $P < 0.05$ and $|\log_2FC| > 1$ (fold change > 2) were considered significant. Variance-stabilizing transformation (VST) was applied to normalized counts for principal component analysis and heatmap visualization.

In addition to count-based analyses, we calculated reads per kilobase of transcript per million mapped reads (RPKM) for each gene across the full dataset, based on the DESeq2-normalized library sizes. For each gene, RPKM was computed as:

$$\text{RPKM} = \frac{10^9 \times C}{N \times L}$$

where C is the raw read count for that gene, N is the total number of mapped reads in the sample, and L is the gene length in base pairs. RPKM values were used for measuring and visualizing expression profiles of candidate genes (e.g., hatching enzyme genes). Statistical significance for each target gene was re-evaluated using Student's t -tests comparing RPKM values between Atlantic and Baltic samples.

Determination of transglutaminase activity for egg envelope hardening

Transglutaminase responsible for the egg hardening has been reported to be localized in the unfertilized egg envelope (31). Therefore, the incorporation of monodansylcadaverine (MDC) into the unfertilized egg envelope was measured. Unfertilized egg envelopes were isolated from the mature ovaries, homogenized in phosphate-buffered saline (PBS), and the supernatant was discarded. After several washes in PBS, approximately 40 mg of unfertilized eggs were placed in a 1.5 mL tube and incubated in 0.5 mL of buffer containing 50 mM Tris-HCl (pH 8.0), 10mM CaCl₂, 2 mM 2-mercaptoethanol, 0.5 mM MDC (Sigma-Aldrich), and varying concentrations of NaCl at 20°C for 30 min. After centrifugation (1000 rpm for 1 min), the precipitate was washed by 1 mL of TBS-T (Tris-buffered saline containing 0.05% Tween 20) containing 10 mM EDTA and briefly homogenized. This washing step was repeated three times, followed by three washes with 1 mL of 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA. The precipitate was then incubated in 200 µL of 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA, 0.5% SDS, and 0.5 mg/mL Proteinase K at 55°C until completely dissolved. Fluorescence intensity of the solution was measured using a plate reader (Varioskan LUX, ThermoFisher) with excitation at 355 and emission at 550 nm. Protein concentration was determined by absorbance at 280 nm. MDC fluorescence standards (10~500 µM/mL) were used for calibration. Enzyme activity was expressed as nmole/30 min/mg protein. Values are presented as means ± S.E. of six measurements, obtained from two independent experiments using ovaries from three different mature females.

Determination of proteolytic activity of the hatching enzyme

Caseinolytic activity was measured using 500 µL of reaction mixture consisting 50 mM Tris-HCl (pH 8.0), 3.3 mg/mL casein, varying concentrations of NaCl, and the enzyme. The mixture was incubated for 1 h at 30°C. The reaction was stopped by adding 125 µL of 20% perchloric acid, followed by incubation in an ice-cold water bath for 10 min. The mixture was then centrifuged at 18.500 g for 10 min at 4°C, and the absorbance of the supernatant at 280 nm (A₂₈₀) was measured.

Estimation of protein content per egg envelope

Forty isolated fertilized egg envelopes were placed in 1.5 mL tubes. They were washed by 1mL of TBS-T (Tris buffered saline-0.05% Tween 20) with brief homogenization to remove cell debris, yolk protein and adhesive layer. This step was repeated three times, then the supernatant was replaced to 1 mL of 50 mM Tris-HCl (pH 8.0) and 0.15 M NaCl. After centrifugation (1000rpm for 1min), the precipitate was incubated in 200 µL of 50 mM Tris-HCl (pH 8.0), 10mM EDTA, 1% SDS and 0.5 mg ProteinaseK at 55°C until the precipitate was completely dissolved. The protein content was determined by absorbance at 280 nm. The values and S.E.M. were expressed from the average of six measurements each in the Baltic and Atlantic herring.

Supplementary Text

Results of proteomic analysis of sperm and egg and modelling of LRRC8 structures

Introduction to volume-regulated anion channels (VRACs) LRRC8C

Volume-regulated anion channels (VRACs), also known as volume-sensitive outwardly rectifying anion channels, play a crucial role in responding to hypotonic stress and maintaining physiological osmotic balance, thereby preventing cell swelling or lysis (54-56). VRACs usually assemble as hexamers of leucine-rich repeat-containing family 8 (LRRC8) proteins, each comprised of a pore domain (PD)—facilitating osmolyte flux—and a leucine-rich repeat domain (LRRD), which likely interacts with the cytoskeleton to relay mechanosensory signals. Mammals possess five paralogues (LRRC8A-E), each conferring different permeability or selectivity. Cryo-EM structures have revealed that mammalian LRRC8A and LRRC8C channels can exhibit hexameric or heptameric stoichiometry, respectively. The LRRC8C orthologue of the Atlantic herring exists as two different paralogues, referred to as C1 (on chromosome 10) and C2 (on chromosome 2). Importantly, C1 is hypothesized to be a house-keeping LRRC8 isoform, whereas C2 is thought to be sperm-specific.

Results

Sperm and egg proteome in Atlantic herring

Mass spectrometric (MS) analysis of three sperm samples identified on average ~3970 proteins (Dataset S1), at a depth comparable to a published sperm proteome from human (6). Notably, 82% of these proteins were detected in all three sperm samples (Dataset S1). The MS analysis of three oocytes yielded on average 5830 proteins, with 81% percent shared across all oocyte samples (Dataset S1).

For the sperm data, the analysis specifically addressed whether LRRC8C2 (Chr2) (A0A6P3VMJ8) is an isoform exclusively expressed in sperm and whether LRRC8C1 (Chr10) (A0A6P8G6J1) is confined to somatic tissues. To resolve LRRC8C2 isoforms, the following three amino-acid sequences of LRRC8C2 (Chr2) alleles (Intermediate, Baltic North, Baltic South) were added to the UniProt database of *C. harengus*:

>lrrc8c_intermediate

MIAVSELWQFTEQPPAFRVLKPWWDFVMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS
ACPLTSNQTGIFPDQTVAAALPTVSSASSGSPDLELKGVTNNLDLQQYKFQINQLCYQTSLSH
WFPKCFPYLVLIQTIIIFMVCSNFWFKFPGSSSKIEHFISILGKCFDSPWTTTRALSEVSGE
HPEEAENLASPEACTQPLLSMRDKTSEEVLDKPVAVASVLDSECEQAKALFERIKKFRV
HVEEGNLLYKTYAYQTILKVFMFIFIIAYNSALVTKIRVLVTCNMDLQDMTGYRHFCIF
TMAHLFSKLADSYLFFVVMYGLISLYTTYWLFHRSLSKEYSFEYVRKETGFRDIPDVKNDF
AFMLHMMDQYDPLYSKRFVFLSEVSENKLIQLNLNHAWTAEKLRRRLQTNSSKRLELQL
LMLPGLPDTIFQLSELQSLKLELIRNATIPAAITQLEELEELSLNQCSLKVNMESISFLK
ANLKVLRVRFVDANELPYWIYVLRNLEELHLIGSLSPSPSKNITLKSRLDLKSLKTLKSLK
SNVTRIPQAIVDLSSRLQQLLIYNDGTKLMTLNSLKKMVNLTELELVHCDLECIPSAIFS
LVNLQLLDLRENKLCSEIEIVSFQHLPKFTCLKLWHNSIAQIPEHIKKLVNLERLYFSHN
KIETLPFHLFLCNKLCYLDLSHNDIGLIPSEVRLHSLQYLSVTCNKIKALPDELFS CRT
LKTLLKLGNNLLSGLSPKIAKLVLLTDLELKGNYFVFLPCELGACQSLKRSGLVVEEALFE
TLPLDVRERMNAE

>lrrc8c_baltic_south

MIAVSELWQFTEQPPAFRVLKPWWDFVMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS
ACPLTSNQTGIFPDQTVAAALPTVSSASSGSPDLELKGVTNNLDLQQYKFQINQLCYQTSLSH
WFPKCFPYLVLIQTIIIFMVCSNFWFKFPGSSSKIEHFISILGKCFDSPWTTTRALSEVSGE

HPEEAENLASPEACTQPLLSMRDKTSEEVLDPKPVAVASVLDKECEQAKALFERIKKFRV
HVEEGNLLYKTYAYQTILKVFMFIFIIAYNSALVTKIRVLVTCNMDLQDMTGYRHFCFIF
TMAHLFSKLADSYLFFVVMYGLISLYTTYWLFHRSLSKEYSFEYVRKETGFRDIPDVKNDF
AFMLHMMDQYDPLYSKRFAVFLSEVSENKLIQLNLNHAHTAEKLRRLQTNSSKRLELQL
LMLPGLPDTIFQLSELQSLKLELIRNATIPAAITQLEEELEELSLNQCSLKVNMEAISFLK
ANLKVLRVRFVDANELPYWIYVLRNLEELHLIGSLSPSPSKNITLKSRLDLKSLKTLKSLK
SNVTRIPQAIVDLSSRLQQLLIYNDGTLKMTLNSLKKMVNLTELELVHCDLECIPSAIFS
LVNLQLLDLRENKLCSEIEIVSFQHLPKFTCLKLWHNSIAQIPEHIKKLVNLERLYFSHN
KIETLPFHLFLCNKLCYLDLSHNDIGLIPSEVRLHLHSLQYLSVTCNKIKALPDELFCRT
LKTLLKLGNNLLSGLSPKIAKLVLLTDLELKGNYFVFLPCELGACQSLKRSGLVVEEALFE
TLPLDVRERMNAE

>lrrc8c_baltic_north

MIAVSELWQFTEQPPAFRVLKFWVEVMDVLSLVMVISVFGCTLQVMEEKMYCLPQRGS
ACPLTSNQTGIFPDQTVAAALPTVSSASSGPDLELKGVTNNLDLQYKFINQLCYQTSLSH
WFPKCFPYLVLIQTIIIFMVCSNFWFKFPGSSSKIEHFISILGKCFDSPWTTTRALSEVSGE
HPEEAENLASPEACTQPLLSMRDKTSEEVLDPKPVAVASVLDKECEQAKALFERIKKFRV
HVEEGNLLYKTYAYQTILKVFMFIFIIAYNSALVTKIRVLVTCNMDLQDMTGYRHFCFIF
TMAHLFSKLADSYLFFVVMYGLISLYTTYWLFHRSLSKEYSFEYVRKETGFRDIPDVKNDF
AFMLHMMDQYDPLYSKRFAVFLSEVSENKLIQLNLNHAHTAEKLRRLQTNSSKRLELQL
LMLPGLPDTIFQLSELQSLKLELIRNATIPAAITQLEELKELSLNQCSLKVNMEAISFLK
ANLKVLRVRFVDANELPYWIYVLRNLEELHLIGSLSPSPSKNITLKSRLDLKSLKTLKSLK
SNVTRIPQAIVDLSSRLQQLLIYNDGTLKMTLNSLKKMVNLTELELVHCDLECIPSAIFS
LVNLQLLDLRENKLCSEIEIVSFQHLPKFTCLKLWHNSIAQIPEHIKKLVNLERLYFSHN
KIETLPFHLFLCNKLCYLDLSHNDIGLIPSEVRLHLHSLQYLSVTCNKIKALPDELFCRT
LKTLLKLGNNLLSGLSPKIAKLVLLTDLELKGNYFVFLPCELGACQSLKRSGLVVEEALFE
TLPLDVRERMNAE

The MS data analysis of sperm detected 27 peptides shared among the LRRC8C2 isoforms and four peptides shared between LRRC8C2 and LRRC8C1 (Fig. S2). These four peptides are located in a protein region that is 100% identical in LRRC8C2 isoforms and LRRC8C1. Importantly, no unique (proteotypic) peptide of LRRC8C1 was detected in sperm. Considering that proteotypic LRRC8C1 peptides were readily detected in four different herring tissues (Dataset S1), we conclude that LRRC8C1 is not present in sperm.

The MS analysis of sperm detected two unique (proteotypic) peptides from the LRRC8C2 Baltic North allele (Fig. S2) arising from the E460K substitution. Consistent with their higher MS1 intensity, the LRRC8C2 Baltic North allele is ranked highest among the other isoforms, supporting the conclusion that it is the predominant LRRC8C2 allele in the sperm samples, consistent with the fact that the Baltic herring for MS analysis was sampled in the Bothnian Sea.

Overlay of cryo-EM structures and AlphaFold2 predictions of monomeric LRRC8 isoforms

Monomeric structures of mammalian LRRC8A and C predicted by AlphaFold2 (57) were compared to their respective cryo-EM structures (58-60). The overlays indicated root-mean-square deviations (RMSD) of 3.22 Å for human LRRC8A and 2.96 Å for mouse LRRC8C. The deviations are likely attributable to missing flexible loop regions in extracellular and cytosolic subdomains in the cryo-EM structures. This satisfactory concordance provides confidence for subsequent AlphaFold predictions of the herring LRRC8C isoforms. AlphaFold2 monomeric models of LRRC8C1 and LRRC8C2 are overlaid in Fig. S4, revealing a minimal aggregate RMSD of 1.17 Å, which indicates that their overall structures are highly similar at the monomer

level. The amino-acid substitutions in herring LRRC8C2, highlighted as yellow spheres, are located in the transmembrane region (TM) and cytoplasmic subdomain (CSD) and may influence the permeation properties of the pore.

Comparison of the ESD selectivity filter in LRRC8A and LRRC8C from mammals and fish

Homo-multimeric LRRC8A channels generate smaller whole-cell currents compared to hetero-multimeric channels containing both LRRC8A and LRRC8C subunits (60), likely due to a narrower selectivity filter in LRRC8A. In LRRC8A, residue R103 within the extracellular subdomain (ESD) projects into the pore lumen, creating a constriction approximately 5 Å in diameter. The six R103 residues at this site collectively provide a high positive charge density, which is thought to electrostatically steer anions through the pore while excluding cations.

By contrast, mouse LRRC8C replaces R103 with the hydrophobic residue L105. This substitution relaxes the selectivity filter in two ways: (i) the smaller side chain of L105 reduces the extent of pore constriction, and (ii) its neutral character lessens coulombic repulsion of cations at the constriction. In herring C1 and C2, positions corresponding to R103 of mammalian LRRC8A are occupied by W107 and L104, respectively - both hydrophobic, similar to L105 in mouse LRRC8C. However, the larger W107 side chain may imitate R103 by conferring a narrower constriction.

To assess the roles of W107 in C1 and L104 in C2, Alphafold2 structures for the respective homo-heptamers were generated. A comparison of the ESD selectivity filters among mammalian LRRC8A, LRRC8C, and herring C1 and C2 (Fig. S4b-e) shows that LRRC8A subunits pack more tightly than LRRC8C subunits, likely due to hexameric vs. heptameric assembly. In LRRC8A hexamers, the R103 side chains penetrate deeply into the pore lumen to create the primary constriction, a feature absent in mammalian LRRC8C and herring C1 and C2, where side chains of L105, W107, and L104 do not project into the pore lumen. In herring C1, neighbouring W107 residues, being ~5 Å apart, are too distant to enable π - π stacking interactions, and their indole rings are oriented away from potential hydrogen-bond acceptors such as carbonyl oxygens of Q108 and N104. Instead, W107 side chains are positioned in the inter-subunit space, likely preventing tight subunit packing through steric effects.

Comparison of electrostatic maps of the ESD selectivity filters (Fig. S3) reveals two main differences: (i) LRRC8A hexameric channels possess a smaller pore diameter than LRRC8C heptameric paralogues, and (ii) LRRC8C paralogues exhibit markedly lower charge density lining the pore entrance. In LRRC8A, residues K98 and D100 form a salt-bridge, stabilizing the tight packing of subunits (Fig. S3a). The K98 side chain projects into the space between subunits, creating pockets of high positive charge around the pore, whereas the D100 side chain runs parallel to the long pore axis, forming an outer ring of negative charge above the pore opening. Deeper into the pore, R103 and H104 together create an inner ring of high positive charge. In LRRC8C paralogues (mammalian, Fig. S3b; herring C1, Fig. S3c; and herring C2, Fig. S3d), the homologous positions to K98, D100, R103, and H104 in human LRRC8A, contain fewer charged amino acids, as detailed in Table S6. Specifically, R103 and H104 in LRRC8A are replaced by L105 and Q106 in mouse LRRC8C, by W107 and Q108 in herring C1, and by L104 and Q105 in herring C2; D100 is replaced by N104 and N101 in herring C1 and C2, respectively. As a result, the inner ring of high positive charge density found in LRRC8A is absent in mammalian LRRC8C and herring C1/C2. Taken together, this shift in charge distribution suggests that homomeric LRRC8C channels are more permissive, allowing the passage of larger and neutral osmolytes such as taurine - unlike the more selective LRRC8A channel.

Comparison of pore radius profiles in homomeric LRRC8A and LRRC8C channels

To understand how structural differences impact permeation and selectivity in LRRC8A versus LRRC8C channels, we analysed pore diameter using the MOLE webserver (61). Pore profiles (Fig. S5) aligned to the long axis of the pore overlaid with the corresponding channel structures, reveal striking differences: mammalian LRRC8C has a short and wide pore (~120 Å length) with a radius ≥ 15 Å for 54% of its length, forming a weakly-prolate spheroid (flattening parameter = -3) (Fig. S5a, b). In contrast, LRRC8A features a long, narrow pore (~160 Å length) with a radius ≥ 15 Å for only 24% of its length, representing a strongly prolate spheroid (flattening parameter = -7). Furthermore, LRRC8A contains multiple constrictions - most prominently at residue R103, but also at T48, P15, and K235 - producing a lobed internal cavity. By comparison, constrictions in LRRC8C are wider (at L105, M48 and P15) or absent (at S222).

The *C. harengus* C1 isoform, likely the house-keeping variant, shows a pore length similar to mammalian LRRC8C (~120 Å length), but its diameter shows hybrid features of both mammalian LRRC8A and LRRC8C. Its radius ≥ 15 Å for only 20% of its length, equating to an intermediate prolate spheroid (flattening parameter = -5). Remarkably, the pore of C2 mirrors mammalian LRRC8A: ~120 Å length, radius ≥ 15 Å for 15% of its length, and a strongly prolate profile (flattening parameter = -7).

Visual examination of pore profiles (Fig. S5e), identifies a constriction at the apex of the C1 pore (W107), similar to mammalian LRRC8C, while the mid-point of the channel (around P15) matches the diameter of mammalian LRRC8A. In C2, no narrow apex constriction is found, resembling mammalian LRRC8C properties. Overall, these findings suggest that C1/C2 channels may support the passage of larger osmolytes due to the lack of a restrictive selectivity filter, paralleling the permeation traits of mammalian LRRC8C.

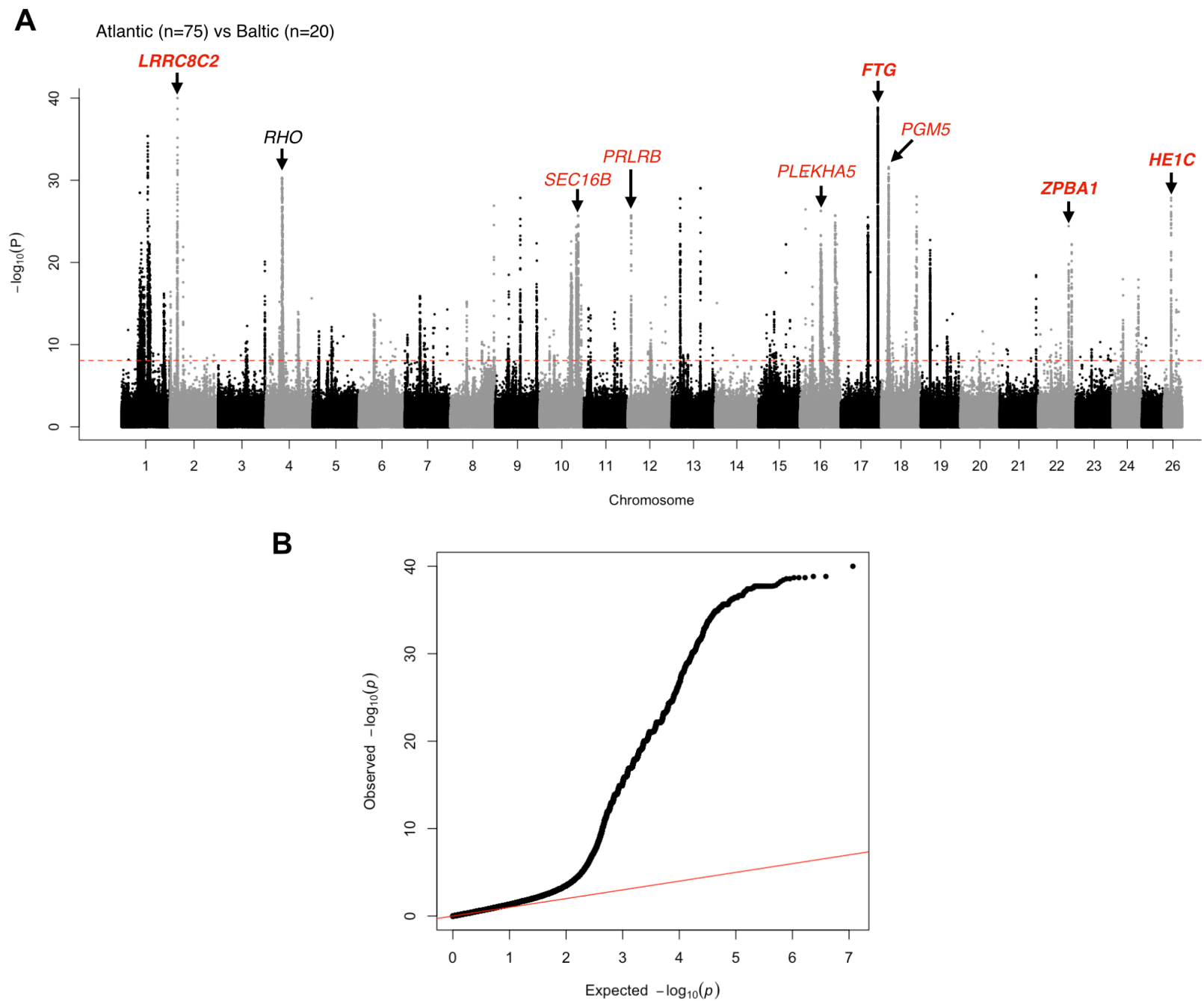
References

1. F. Han *et al.*, Ecological adaptation in Atlantic herring is associated with large shifts in allele frequencies at hundreds of loci. *eLife* **9**, e61076 (2020).
2. J. Goodall *et al.*, Evolution of fast-growing piscivorous herring in the young Baltic Sea. *Nature Communications* **15**, 10707 (2024).
3. S. Purcell *et al.*, PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
4. M. E. Pettersson *et al.*, A long-standing hybrid population between Pacific and Atlantic herring in a subarctic fjord of Norway. *Genome Biol Evol* **15**, evad069 (2023).
5. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528 (2018).
6. E. Grahn *et al.*, Control of intracellular pH and bicarbonate by CO₂ diffusion into human sperm. *Nature Communications* **14**, 5395 (2023).
7. C. S. Hughes *et al.*, Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* **14**, 68-85 (2019).
8. C. S. Hughes *et al.*, Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* **10**, 757 (2014).
9. F. Meier *et al.*, diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods* **17**, 1229-1236 (2020).
10. R. Bruderer *et al.*, Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol Cell Proteomics* **14**, 1400-1410 (2015).
11. S. Tyanova, T. Temu, J. Cox, The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301-2319 (2016).

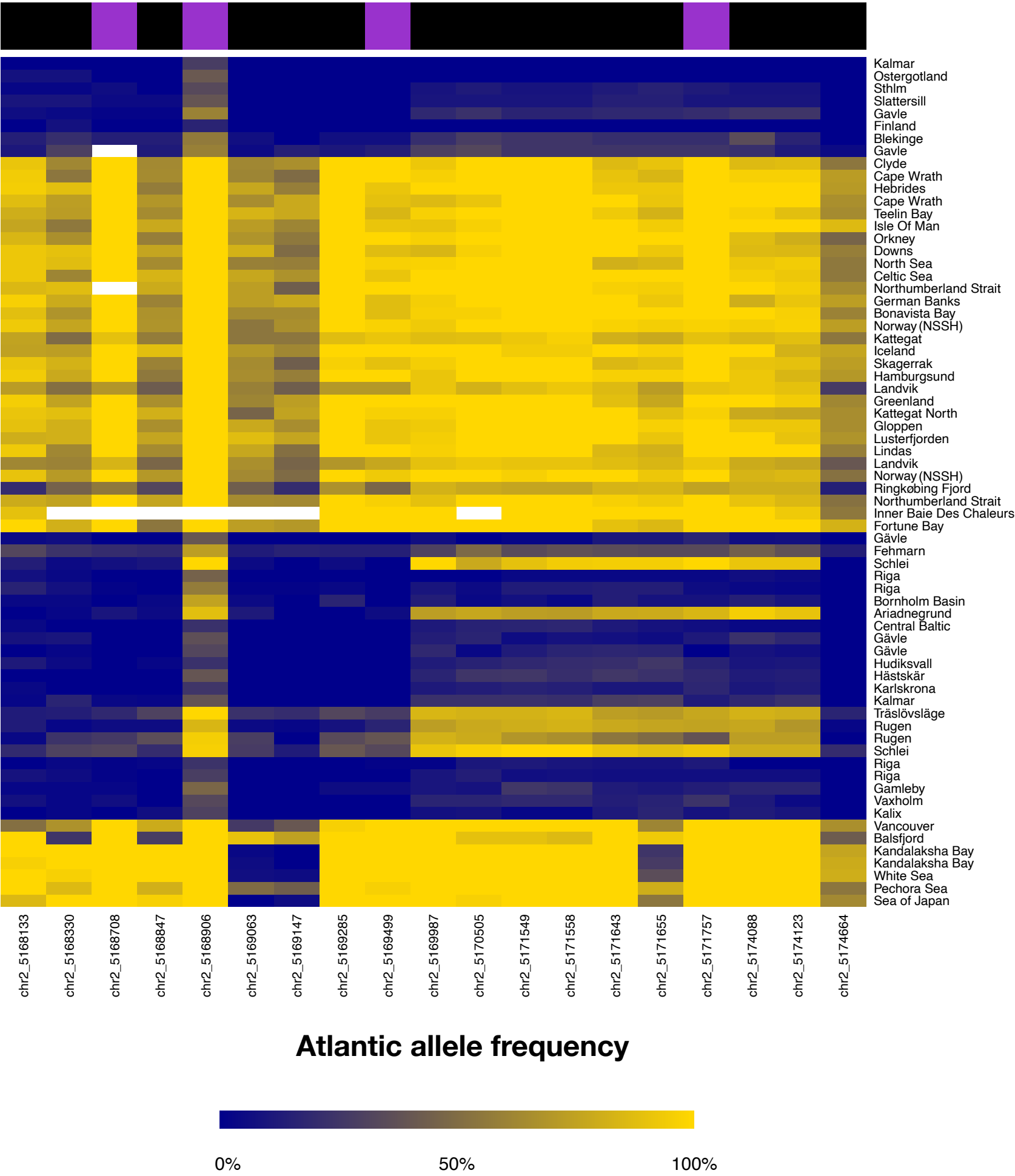
12. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).
13. J. Cox *et al.*, Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**, 1794-1805 (2011).
14. B. Schwanhäusser *et al.*, Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011).
15. M. Jamsandekar *et al.*, The origin and maintenance of supergenes contributing to ecological adaptation in Atlantic herring. *Nature Communications* **15**, 9136 (2024).
16. M. Jamsandekar *et al.*, Unheralded high MHC Class II polymorphism in the abundant Atlantic herring resolved by long-read sequencing. *bioRxiv* 10.1101/2025.06.08.658498, 2025.2006.2008.658498 (2025).
17. H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170-175 (2021).
18. M. Alonge *et al.*, RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol* **20**, 224 (2019).
19. O. Gotoh, Spaln3: improvement in speed and accuracy of genome mapping and spliced alignment of protein query sequences. *Bioinformatics* **40** (2024).
20. M. Manni, M. R. Berkeley, M. Seppey, F. A. Simão, E. M. Zdobnov, BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution* **38**, 4647-4654 (2021).
21. P. Klemm, P. F. Stadler, M. Lechner, Proteinortho6: pseudo-reciprocal best alignment heuristic for graph-based detection of (co-)orthologs. *Front Bioinform* **3**, 1322477 (2023).
22. B. Buchfink, K. Reuter, H. G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* **18**, 366-368 (2021).
23. J. L. Steenwyk *et al.*, OrthoSNAP: A tree splitting and pruning algorithm for retrieving single-copy orthologs from gene family trees. *PLoS Biol* **20**, e3001827 (2022).
24. Z. Chen *et al.*, De novo assembly of the goldfish (*Carassius auratus*) genome and the evolution of genes after whole-genome duplication. *Sci Adv* **5**, eaav0547 (2019).
25. F. Abascal, R. Zardoya, M. J. Telford, TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* **38**, W7-13 (2010).
26. K. Katoh, D. M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772-780 (2013).
27. J. W. Brown, J. F. Walker, S. A. Smith, Phyx: phylogenetic tools for unix. *Bioinformatics* **33**, 1886-1888 (2017).
28. Z. Yang, PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591 (2007).
29. J. Huerta-Cepas, F. Serra, P. Bork, ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol* **33**, 1635-1638 (2016).
30. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T.-Y. Lam, ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28-36 (2017).
31. S. Yasumasu *et al.*, Transglutaminase mediates the hardening of fish egg envelope produced by duplication of factor XIIIa gene during the evolution of Teleostei. *J Biochem* **176**, 427-436 (2024).

32. M. Suyama, D. Torrents, P. Bork, PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* **34**, W609-W612 (2006).
33. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587-589 (2017).
34. B. Q. Minh *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534 (2020).
35. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518-522 (2017).
36. A. Shumate, S. L. Salzberg, Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639-1643 (2021).
37. P. G., P. M., GFF Utilities: GffRead and GffCompare. *FL1000Research* **9**, 304 (2020).
38. I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**, W78-W82 (2024).
39. E. Paradis, pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419-420 (2010).
40. M. Kawaguchi *et al.*, Intron-loss evolution of hatching enzyme genes in Teleostei. *BMC Evol Biol* **10**, 260 (2010).
41. J. T. Robinson *et al.*, Integrative Genomics Viewer. *Nat. Biotech.* **29**, 24-26 (2011).
42. T. Nagasawa, M. Kawaguchi, K. Nishi, S. Yasumasu, Molecular evolution of hatching enzymes and their paralogous genes in vertebrates. *BMC Ecology and Evolution* **22**, 9 (2022).
43. G. Marçais *et al.*, MUMmer4: A fast and versatile genome alignment system. *PLoS Comput Biol* **14**, e1005944 (2018).
44. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410 (1990).
45. C. Camacho *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
46. H. T., A. M., v. A. B., W. D., H. K., gggenomes: effective and versatile visualizations for comparative genomics. *arXiv:2411.13556* doi: 10.48550/arXiv.2411.13556 (2024).
47. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
48. F. Berg, A. Slotte, L. Andersson, A. Folkvord, Genetic origin and salinity history influence the reproductive success of Atlantic herring. *Marine Ecology Progress Series* **617-618**, 81-94 (2019).
49. P. Ewels, M. Magnusson, S. Lundin, M. Kaller, MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016).
50. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915 (2019).
51. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
52. M. Pertea *et al.*, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology* **33**, 290-295 (2015).

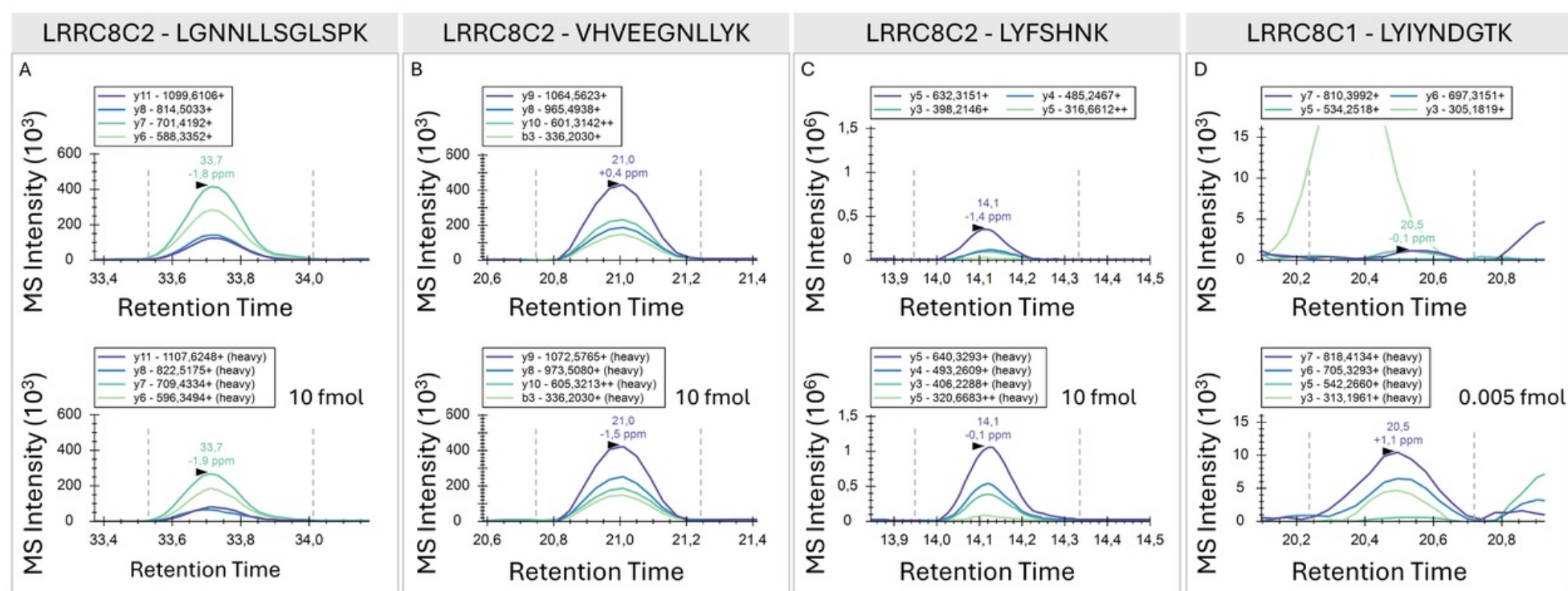
53. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
54. T. J. Jentsch, VRACs and other ion channels and transporters in the regulation of cell volume and beyond. *Nat Rev Mol Cell Biol* **17**, 293-307 (2016).
55. E. K. Hoffmann, I. H. Lambert, S. F. Pedersen, Physiology of cell volume regulation in vertebrates. *Physiol Rev* **89**, 193-277 (2009).
56. J. Osei-Owusu, J. Yang, M. D. C. Vitery, Z. Qiu, Molecular Biology and Physiology of Volume-Regulated Anion Channel (VRAC). *Curr Top Membr* **81**, 177-203 (2018).
57. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
58. D. Deneka, M. Sawicka, A. K. M. Lam, C. Paulino, R. Dutzler, Structure of a volume-regulated anion channel of the LRRC8 family. *Nature* **558**, 254-259 (2018).
59. G. Kasuya *et al.*, Cryo-EM structures of the human volume-regulated anion channel LRRC8. *Nature Structural & Molecular Biology* **25**, 797-804 (2018).
60. S. Rutz, D. Deneka, A. Dittmann, M. Sawicka, R. Dutzler, Structure of a volume-regulated heteromeric LRRC8A/C channel. *Nat Struct Mol Biol* **30**, 52-61 (2023).
61. L. Pravda *et al.*, MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res* **46**, W368-w373 (2018).



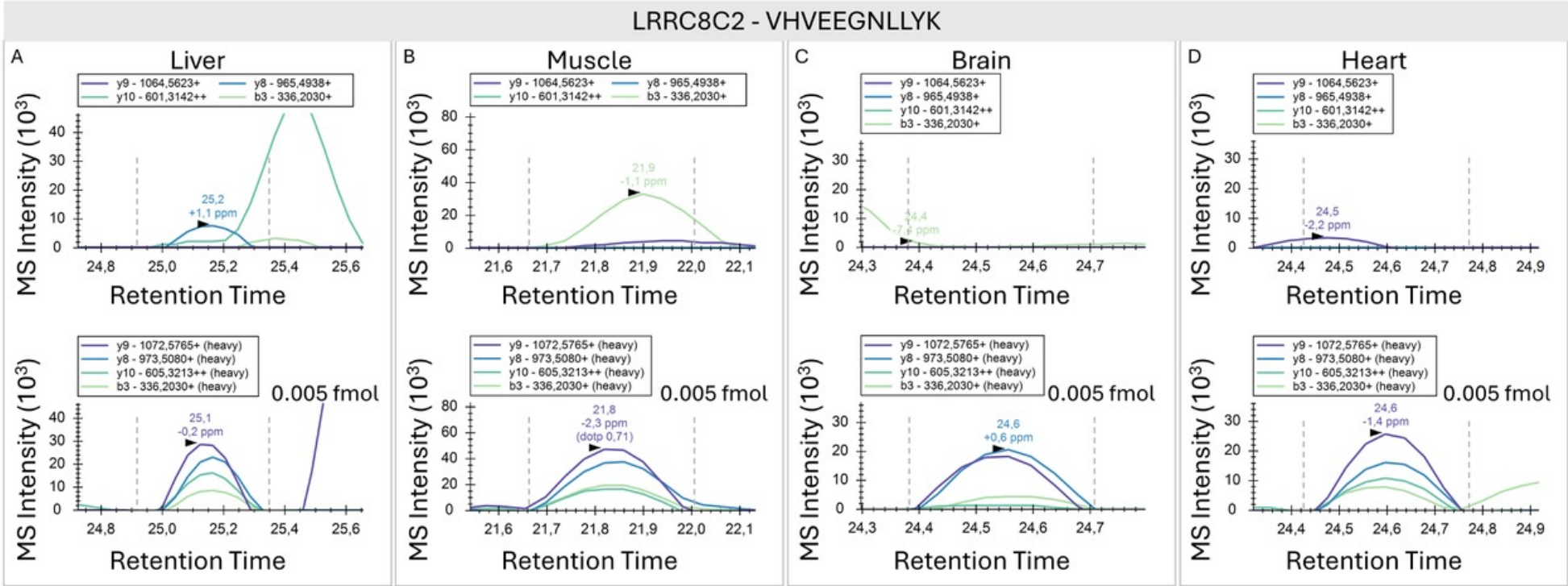
Supplementary Fig. 1. Genome-wide association analysis between Atlantic and Baltic herring populations. (A) Manhattan plot showing results obtained using logistic regression implemented in PLINK, treating Baltic individuals ($n = 20$) as cases and Atlantic individuals ($n = 75$) as controls. Genomic control correction was applied to adjust for potential residual population stratification. The y-axis represents $-\log_{10}(P\text{-values})$. The dashed red line indicates the Bonferroni-corrected genome-wide significance threshold ($P < 8.1 \times 10^{-9}$). Highlighted loci correspond to candidate genes located within the strongest association peaks. (B) Quantile–quantile (QQ) plot showing the distribution of observed versus expected $-\log_{10}(P\text{-values})$, used to assess overall calibration of test statistics and deviation from the null expectation.



Supplementary Fig. 2. Allele frequency heatmap showing highly differentiated SNPs at the *LRRC8C2* locus in a complete set of population samples from Pacific, Atlantic and Baltic herring. Purple boxes above the heatmap indicate missense positions. Note that *LRRC8C2* is encoded on the negative strand, reversing the SNP order.



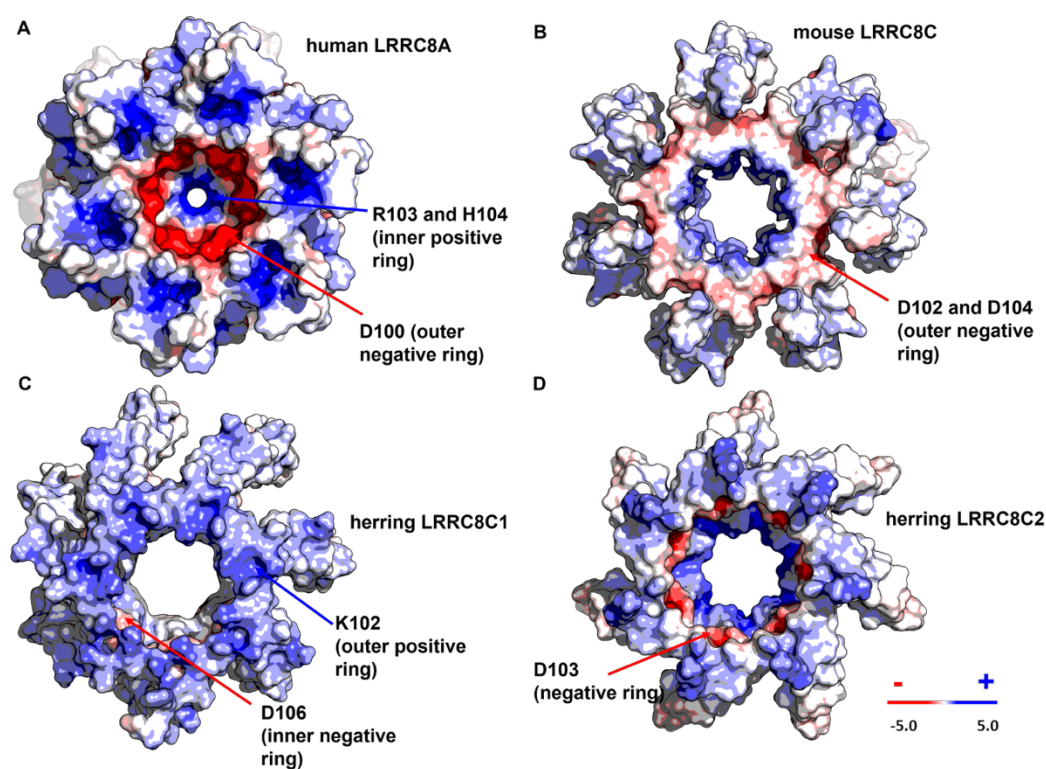
Supplementary Fig. 3. PRM LC-MS/MS analysis of three of the four monitored LRRC8C2 peptides in herring sperm. These LRRC8C2 peptides were readily detected, whereas the LRRC8C1 peptide LYIYNDGTK was not detected. The panels show exemplary PRM transition channels with retention time on the x-axis and MS signal intensity on the y-axis. **(A)** MS signal intensity of endogenous (upper panel) and heavy (lower panel) LGNNLLSGLSPK peptide corresponding to LRRC8C2, in herring sperm peptide background; 10 fmol of heavy peptide were loaded on column. **(B)** MS signal intensity of endogenous (upper panel) and heavy (lower panel) VHVEEGNLLYK peptide corresponding to LRRC8C2 in herring sperm peptide background; 10 fmol of heavy peptide were loaded on the column. **(C)** MS signal intensity of endogenous (upper panel) and heavy (lower panel) LYFSHNK peptide corresponding to LRRC8C2 in herring sperm peptide background; 10 fmol of heavy peptide were loaded on column. **(D)** MS signal intensity of endogenous (upper panel) and heavy (lower panel) LYIYNDGTK peptide corresponding to LRRC8C1 in herring sperm peptide background; 0.005 fmol of heavy peptide were loaded on the column. Endogenous peptide was not detected.



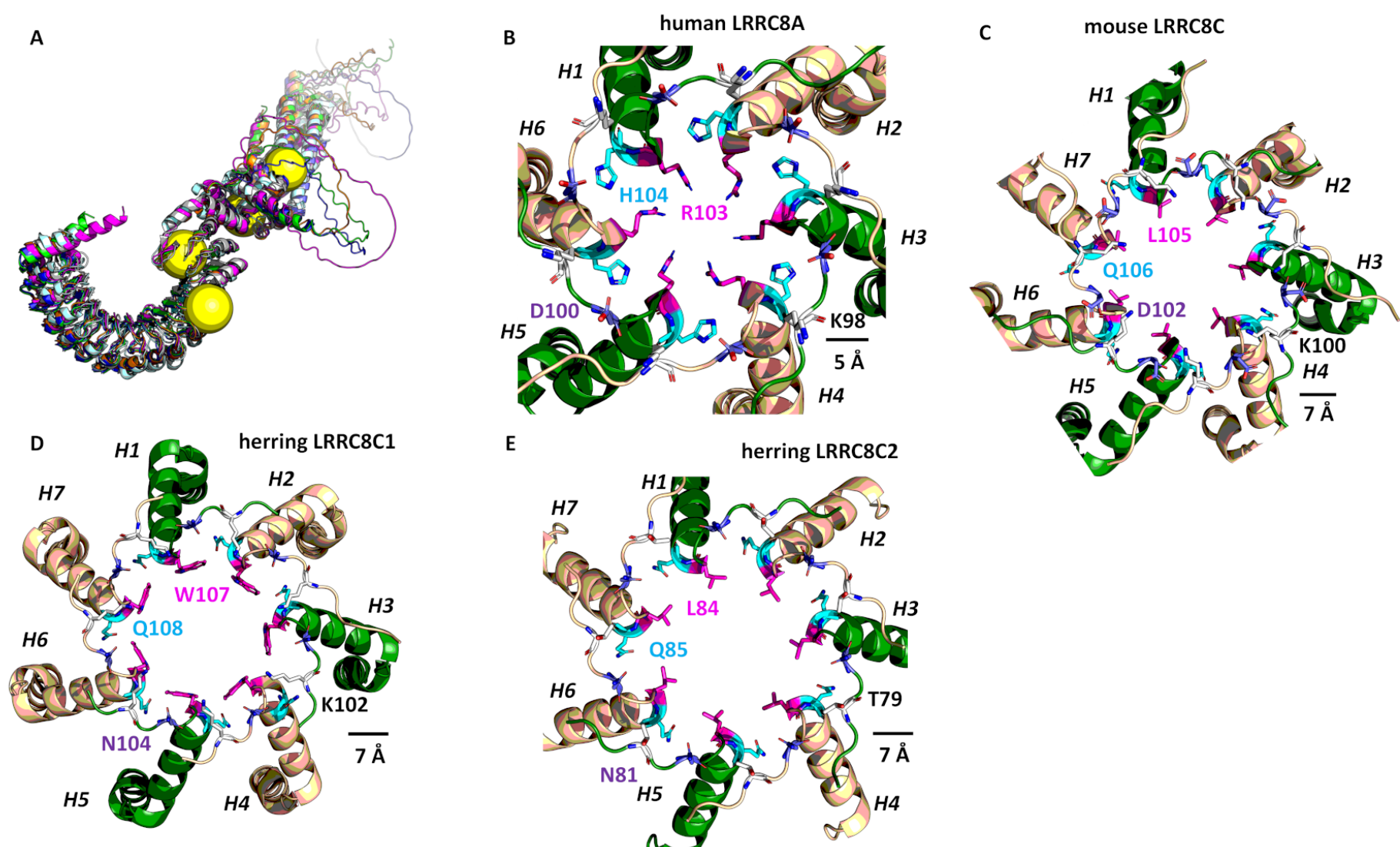
Supplementary Fig. 4. PRM LC–MS/MS analysis of the LRRC8C2 peptide VHVEEGNLLYK in liver, muscle, brain, and heart. Exemplary transition channels from the most sensitive measurements are shown, with retention time on the x-axis and MS signal intensity on the y-axis. Each panel shows MS signal intensities of endogenous (upper panel) and heavy (lower panel) VHVEEGNLLYK peptide in the respective tissue peptide background. For all tissue samples, the lowest determined LOD was 0.005 fmol per μg of protein. Endogenous peptide was not detected in any tissue. (A) Liver. (B) Muscle. (C) Brain. (D) Heart.

Atlantic (A0A6P3VMJ8)	MIAVSELWQFTEQPPAFRVLKPWW	DVFMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS	60	
Intermediate	MIAVSELWQFTEQPPAFRVLKPWW	DVFMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS	60	
BalticN	MIAVSELWQFTEQPPAFRVLKPWW	EVFMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS	60	
BalticS	MIAVSELWQFTEQPPAFRVLKPWW	EVFMDVLSLVMLVISVFGCTLQVMEEKMYCLPQRGS	60	
LRR8C1 (A0A6P8G6J1)	MIPVTEFRQFSEQQPAFRVLKPWW	DVFTDYL SVIMLMIGVFGCTLQVMQDKIICLPQGNL	60	
	** *: *: *: * *****: * * *: *: *: *****: *: * * *			
Atlantic (A0A6P3VMJ8)	ACPLTSNQTGIFPDQTV----	AALPTVSSASSGPSDELEKGV	TNNLDLQQYKFINQLCYQ	116
Intermediate	ACPLTSNQTGIFPDQTV----	AALPTVSSASSGPSDELEKGV	TNNLDLQQYKFINQLCYQ	116
BalticN	ACPLTSNQTGIFPDQTV----	AALPTVSSASSGPSDELEKGV	TNNLDLQQYKFINQLCYQ	116
BalticS	ACPLTSNQTGIFPDQTV----	AALPTVSSASSGPSDELEKGV	TNNLDLQQYKFINQLCYQ	116
LRR8C1 (A0A6P8G6J1)	PVPP-TNQTGILPNQSEPEFALP	SPTSVSPAPVVQEMRGLKTNLDWQ	QYSFINQMCYE	119
	* : * * * *: * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *: * * *:			

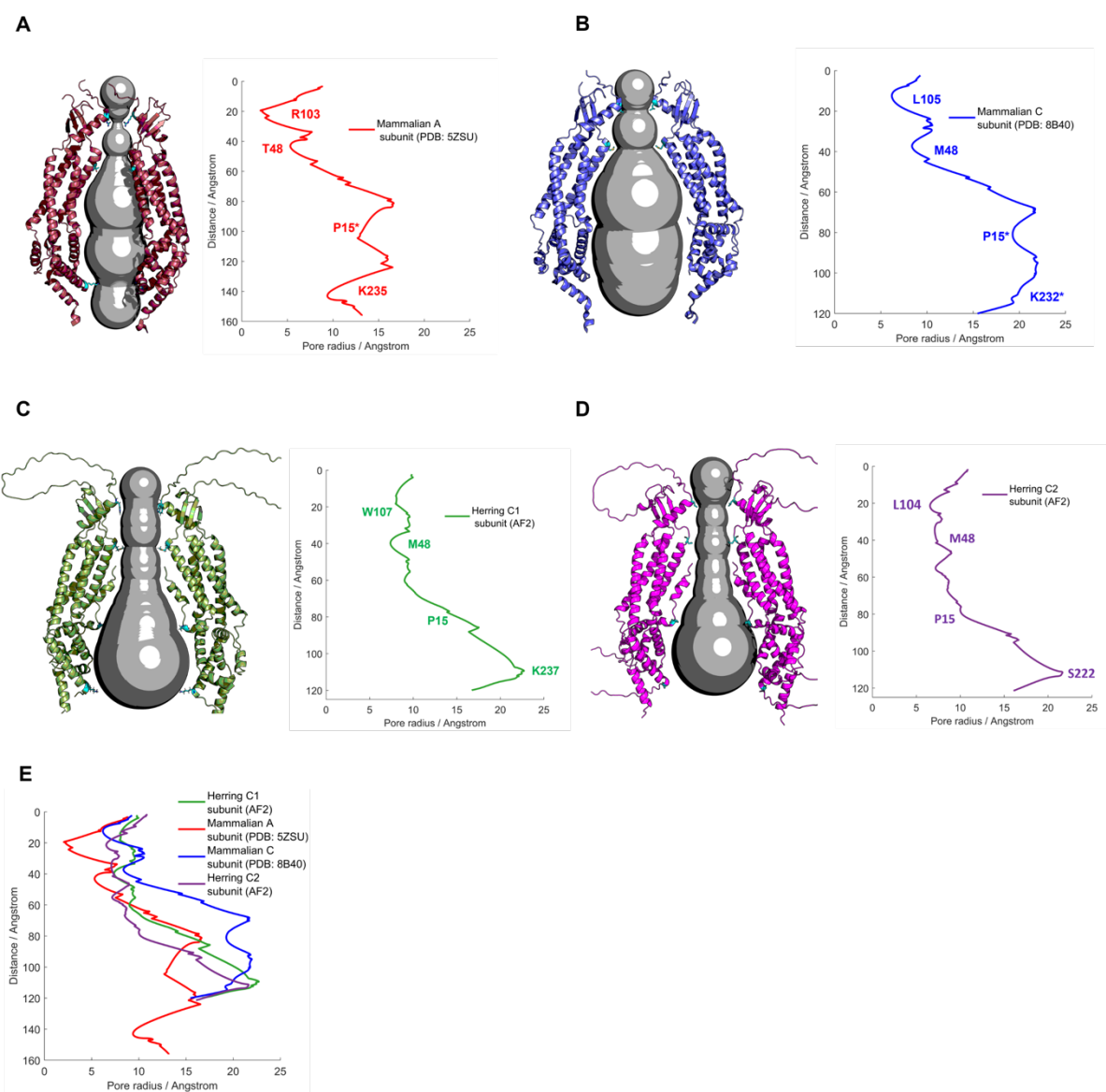
Supplementary Fig. 5. Sequence alignment using ClustalX of the proteins LRR8C2 (Chr2) Atlantic (A0A6P3VMJ8), LRR8C2 Intermediate allele, LRR8C2 Baltic North allele, LRR8C2 Baltic South allele, and LRR8C1 (Chr10) (A0A6P8G6J1). Single amino-acid substitution within the sequences of A0A6P3VMJ8, Intermediate allele, Baltic north, and Baltic south alleles are in red. Peptides of the proteins identified by mass spectrometry in the sperm sample are highlighted in yellow. Two identified, unique peptides for LRR8C2 Baltic north allele, which are not present in the other isoforms, are highlighted in green. Four of the common peptides located in the 100% identical protein region also match to the protein LRR8C1 (Chr10) (UniProt A0A6P8G6J1). No unique peptides for LRR8C1 (Chr10) (A0A6P8G6J1) could be identified which indicates the absence of this isoform in sperm. For detailed information of identified LRR8C2 peptides, see Dataset S1a.



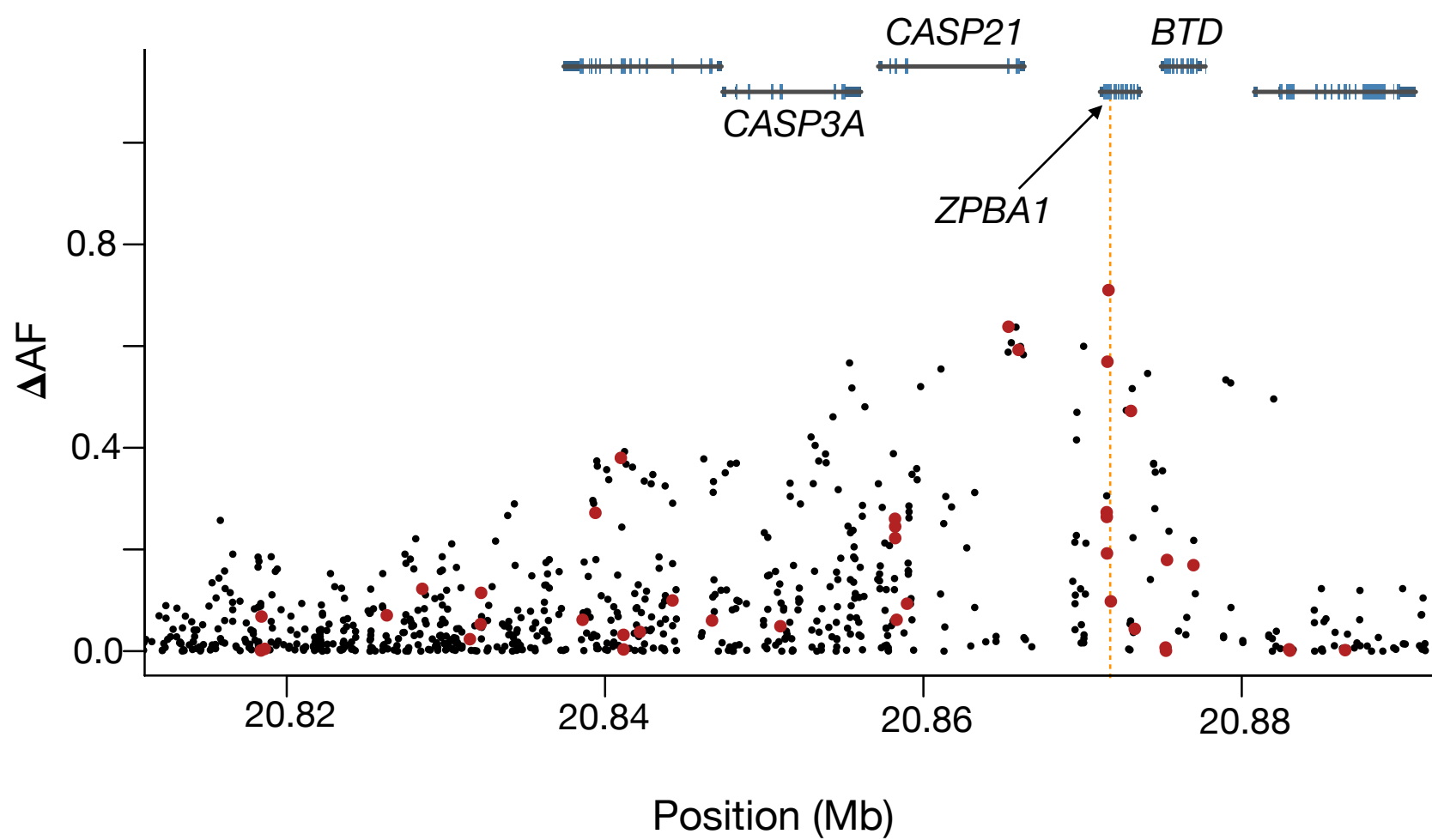
Supplementary Fig. 6. Electrostatic maps of the pore opening of various LRRC8 channels. (A) Surface representation of the ESD of hexameric *H. sapiens* LRRC8A. Arrows indicate the outer ring of negative charges (at the channel's orifice) conferred by residue D100 (red) and the inner ring (within the pore lumen) of positive charge conferred by residues R103 and H104 (blue). (B) Surface representation of the ESD of heptameric *M. musculus* LRRC8C. (C) Surface representation of the ESD of heptameric *C. harengus* LRRC8C1. (D) Surface representation of the ESD of heptameric *C. harengus* LRRC8C2. All electrostatic maps were calculated using the APBS plug-in of Pymol. An arrow indicates the ring of negative charge conferred by residue D103. The charge density scale is indicated (and is the same for (A)-(C)).



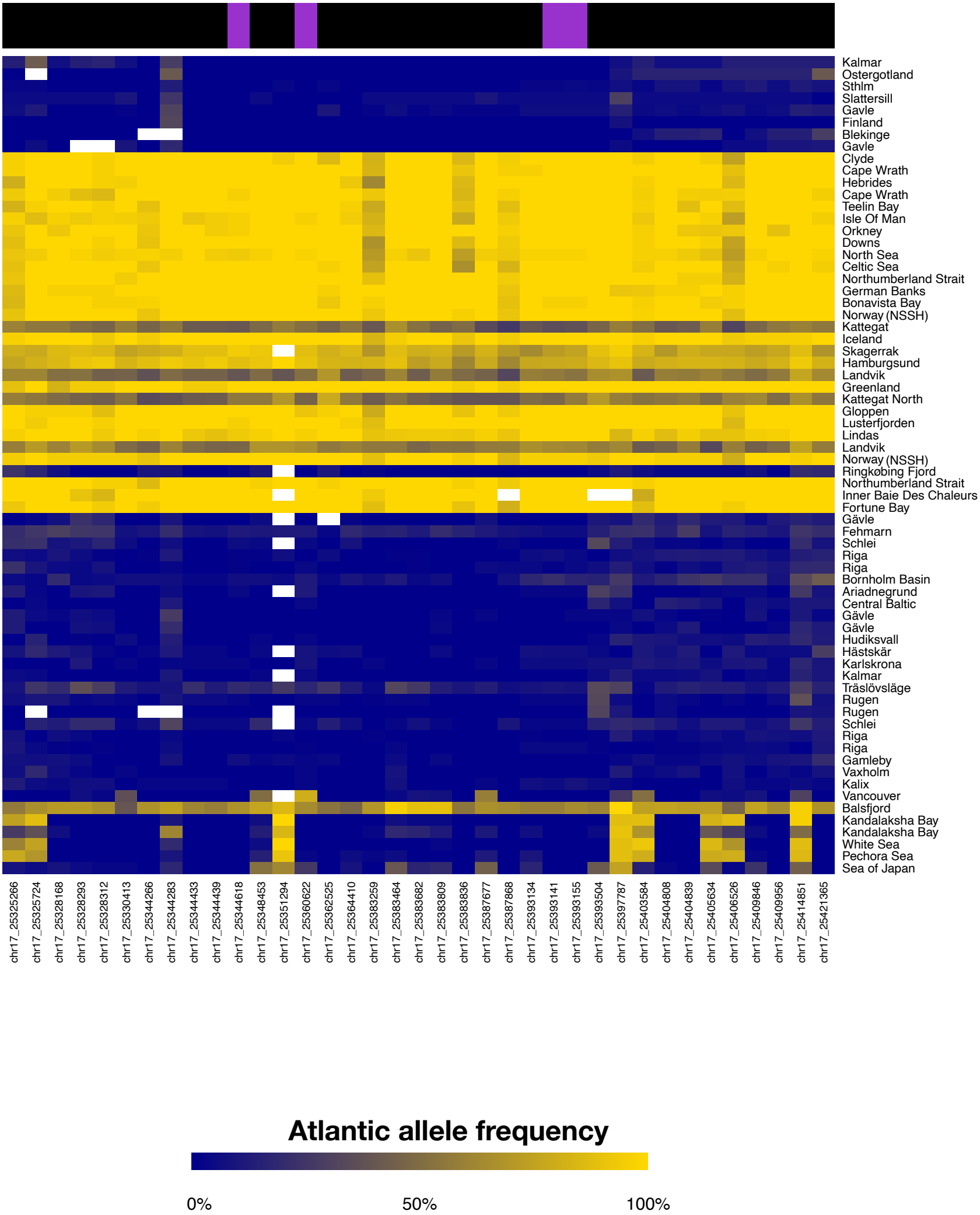
Supplementary Fig. 7. Bird's-eye view of the selectivity filters of various LRRC8 channels. (A) Overlay of monomer structures of LRRC8: *M. musculus* C (PDB: 8B40), *H. sapiens* A (PDB: 5ZSU), *M. musculus* C and *H. sapiens* A (AlphaFold2), *C. harengus* C1 of chromosome 10 and C2 of chromosome 2 (AlphaFold2) shown in grey, cyan, green, magenta, blue, and orange cartoon representation. The residues corresponding to the single-nucleotide polymorphisms (SNPs): D25, V254, K452, and L515 in *C. harengus membras* are shown as yellow spheres. (B) Cartoon of the ESD of hexameric human LRRC8A channel. (C) Cartoon of the ESD of heptameric *M. musculus* LRRC8C channel. (D) Cartoon of the ESD of heptameric *C. harengus* LRRC8C1 subunit. (E) Cartoon of the ESD of heptameric *C. harengus* LRRC8C2 subunit. In all structures, subunits are indicated *H1-6* or *H1-7*, clockwise from centre-top, and residues implicated in anion selectivity (A), or at homologous positions (B)-(E) are labelled, colour-coded, and shown in stick representation.



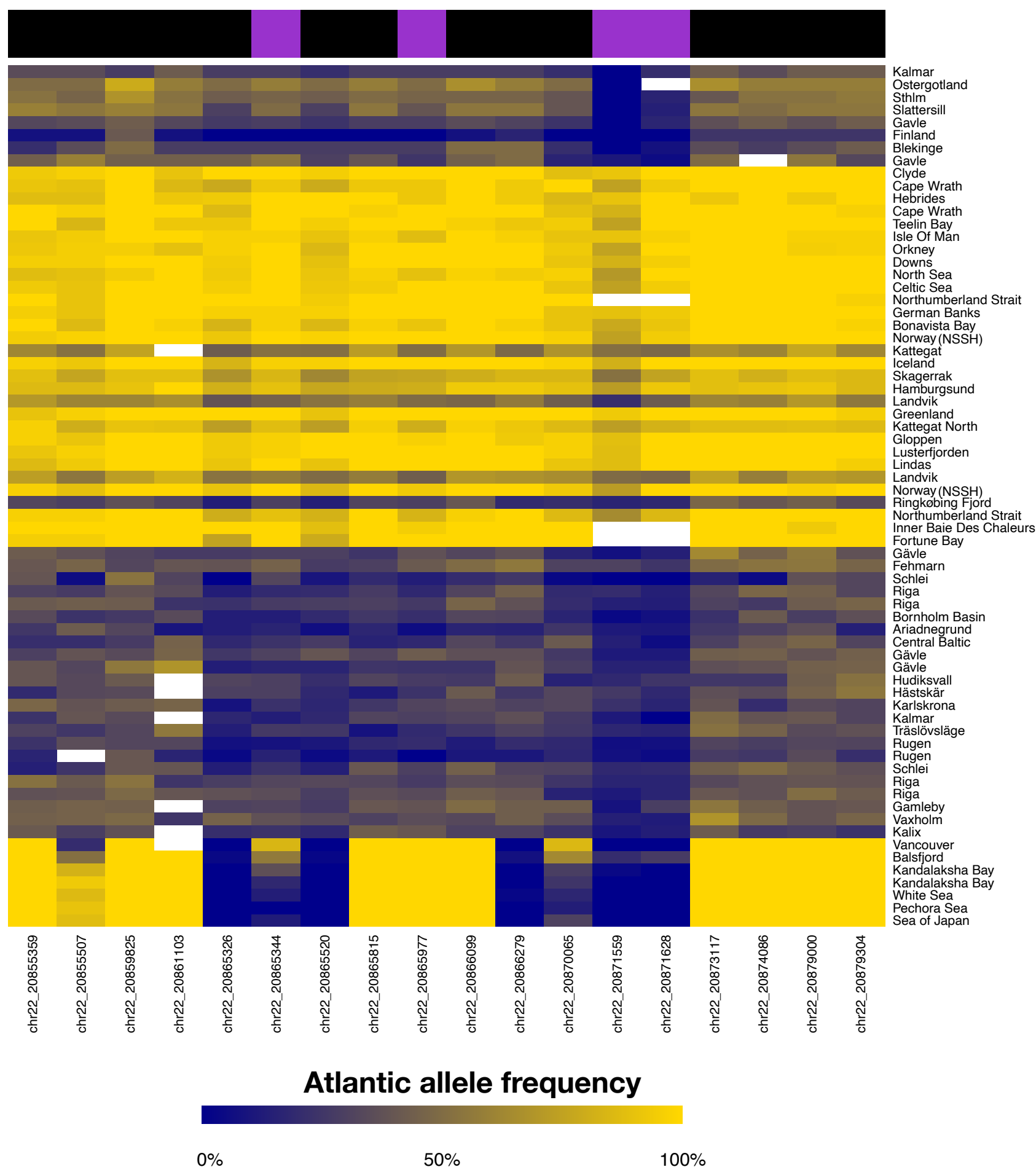
Supplementary Fig. 8. Pore profiles of various LRRC8 channels. (A) (left) Cartoon of hexameric *H. sapiens* LRRC8A (PDB: 5ZSU), with 4 of 6 subunits removed. (right) The pore profile. (B) (left) Cartoon of heptameric *M. musculus* LRRC8C (PDB: 8B40), with 5 of 7 subunits removed. (right) The pore radius profile. (C) Structure of heptameric *C. harengus* LRRC8C1 (AF2) shown in cartoon representation, with 5 of 7 subunits removed. (right) The pore profile. (D) Cartoon of heptameric *C. harengus* LRRC8C2 (AF2), with 5 of 7 subunits removed. In all structures the pores are shown as grey spheres, and residues at homologous positions to P15, T48, R103, and K235 in human LRRC8A are shown in stick representation. (right) The pore profile. All pore profiles were calculated using the MOLE online webserver (mole.upol.cz), and the approximate positions of residues at each constriction are labelled. Residues missing from electron density maps are marked with an asterisk and the position was estimated by aligning an AF2 structure of monomeric LRRC8 with a subunit of each homo-multimer. (E) An overlay of all pore profiles for comparison.



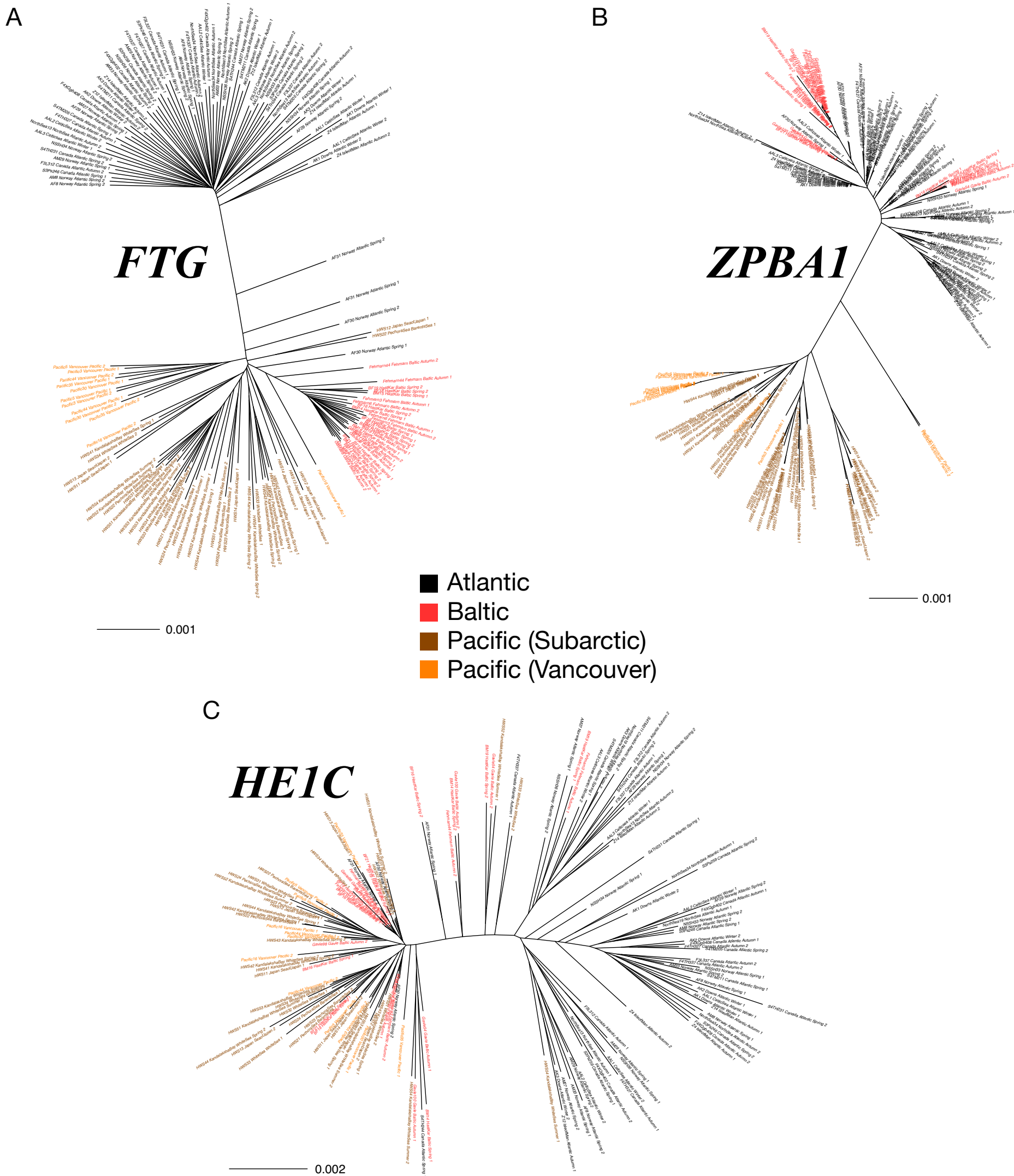
Supplementary Fig. 9. Allele frequency contrasts comparing Atlantic spring-spawners against Baltic spring-spawners for the region on chromosome 22 harbouring *ZPBA1*. SNPs marked in red represent missense mutations.



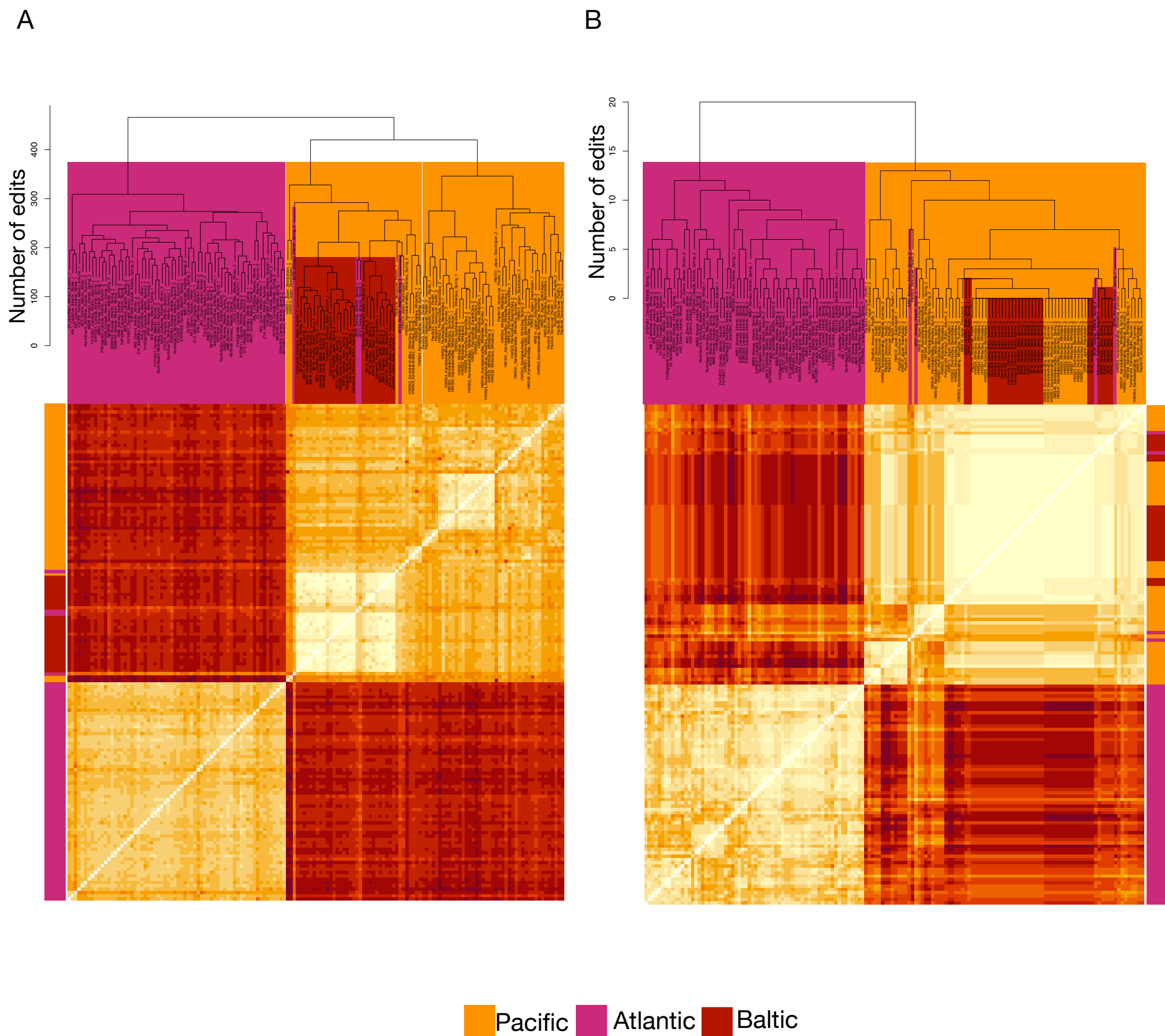
Supplementary Fig. 10. Allele frequency heatmap showing highly differentiated SNPs at the *FTG* locus on chromosome 17 in a complete set of population samples of Pacific, Atlantic and Baltic herring. Purple boxes in the top row indicate missense positions.



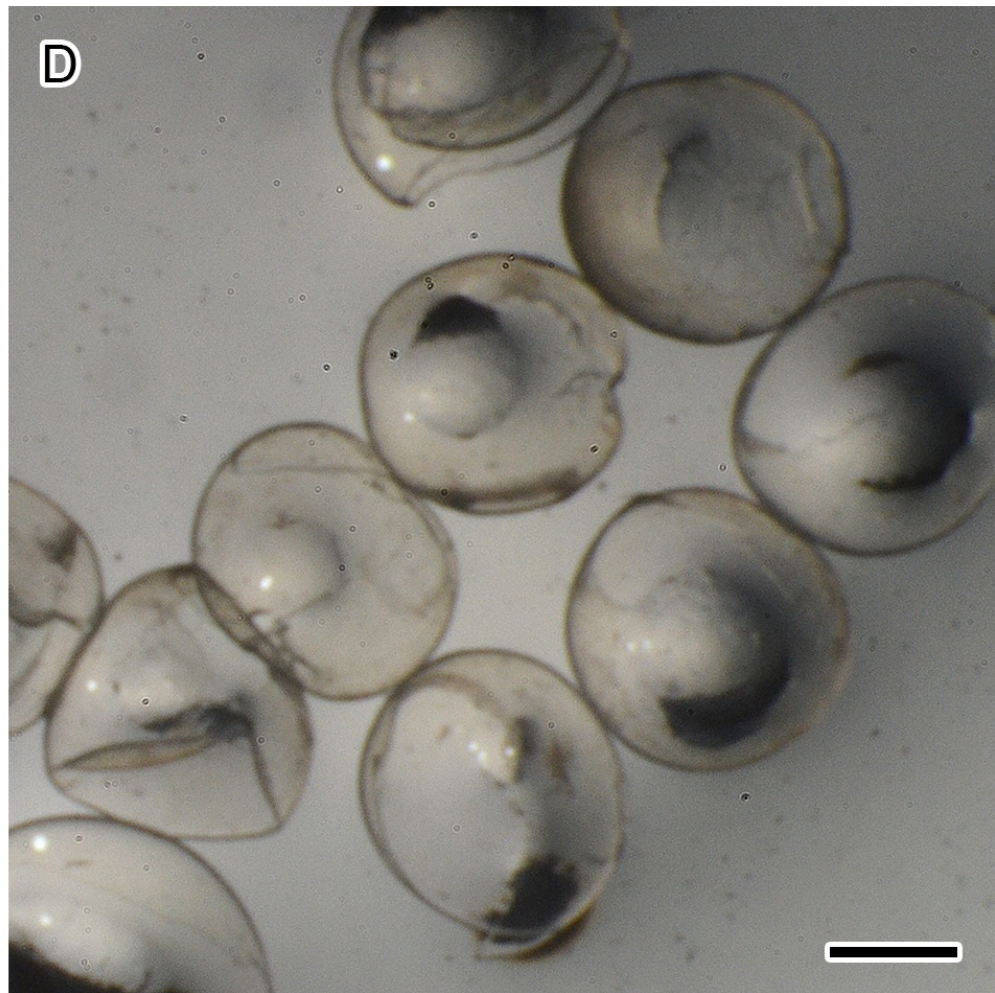
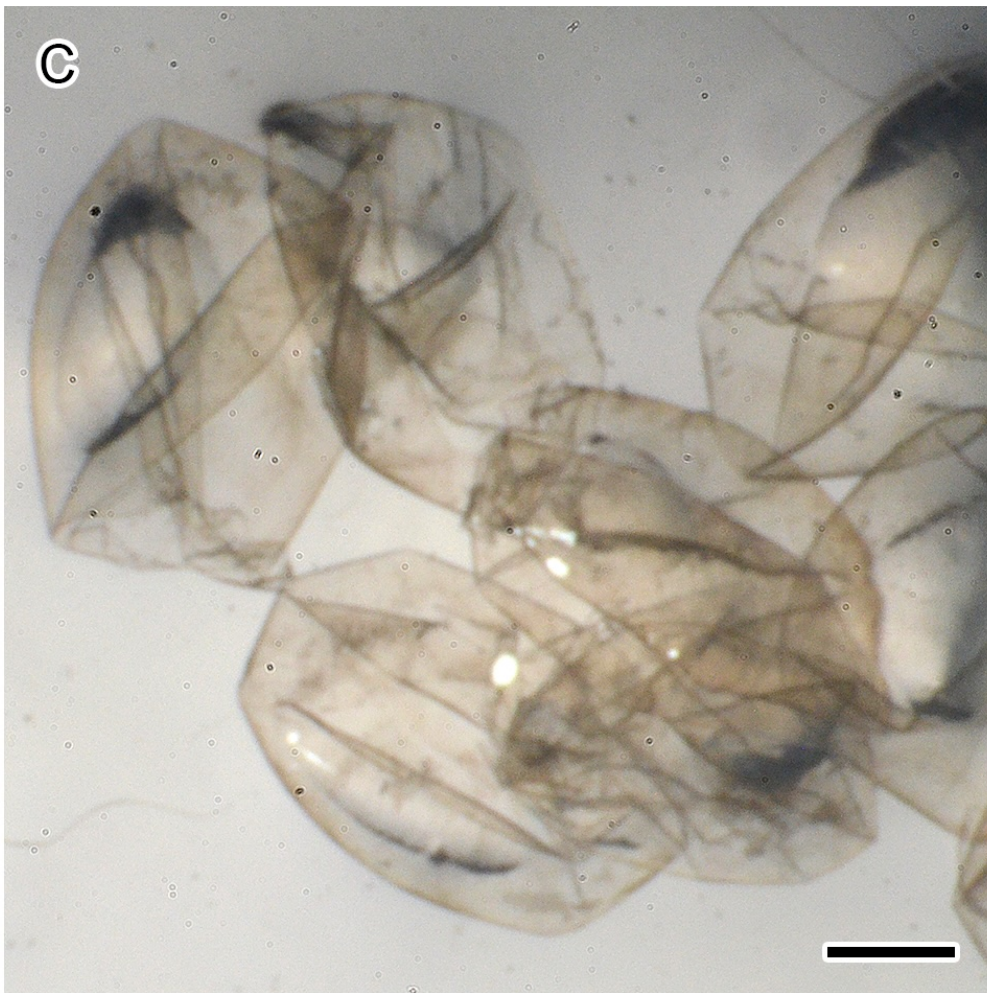
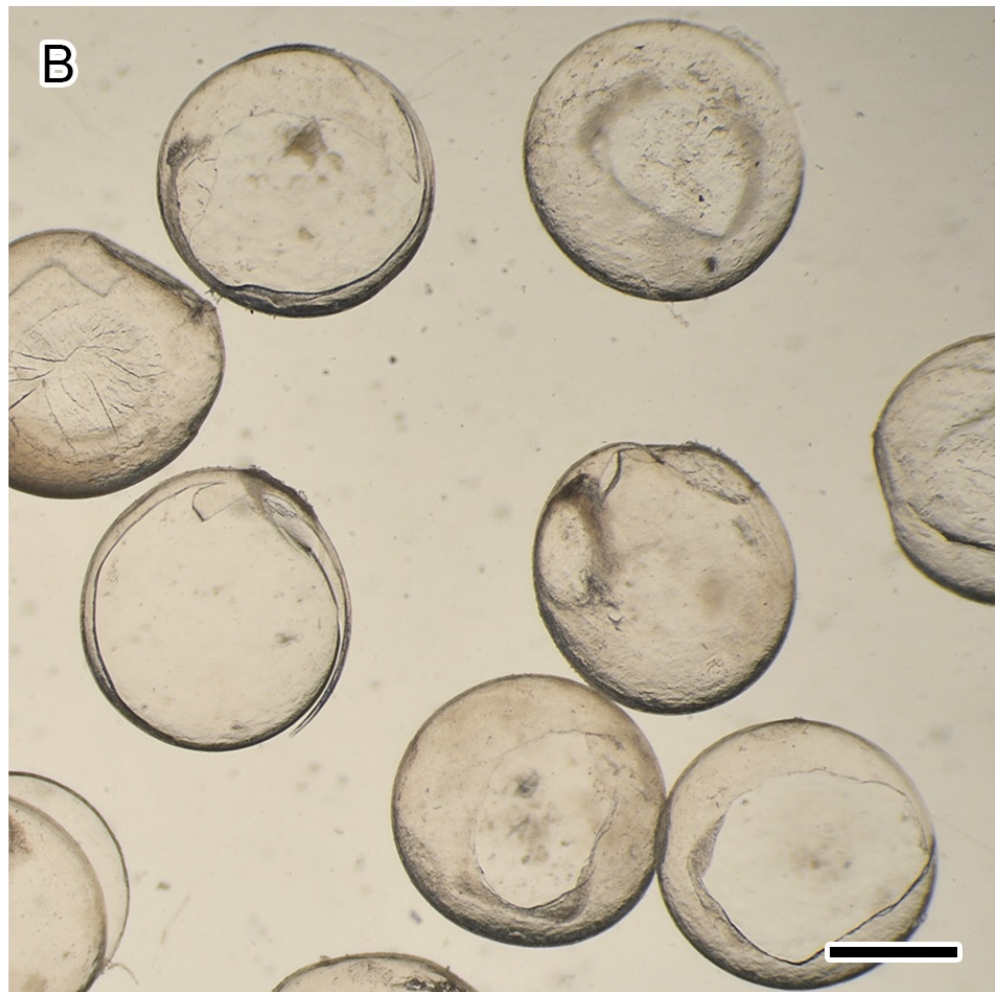
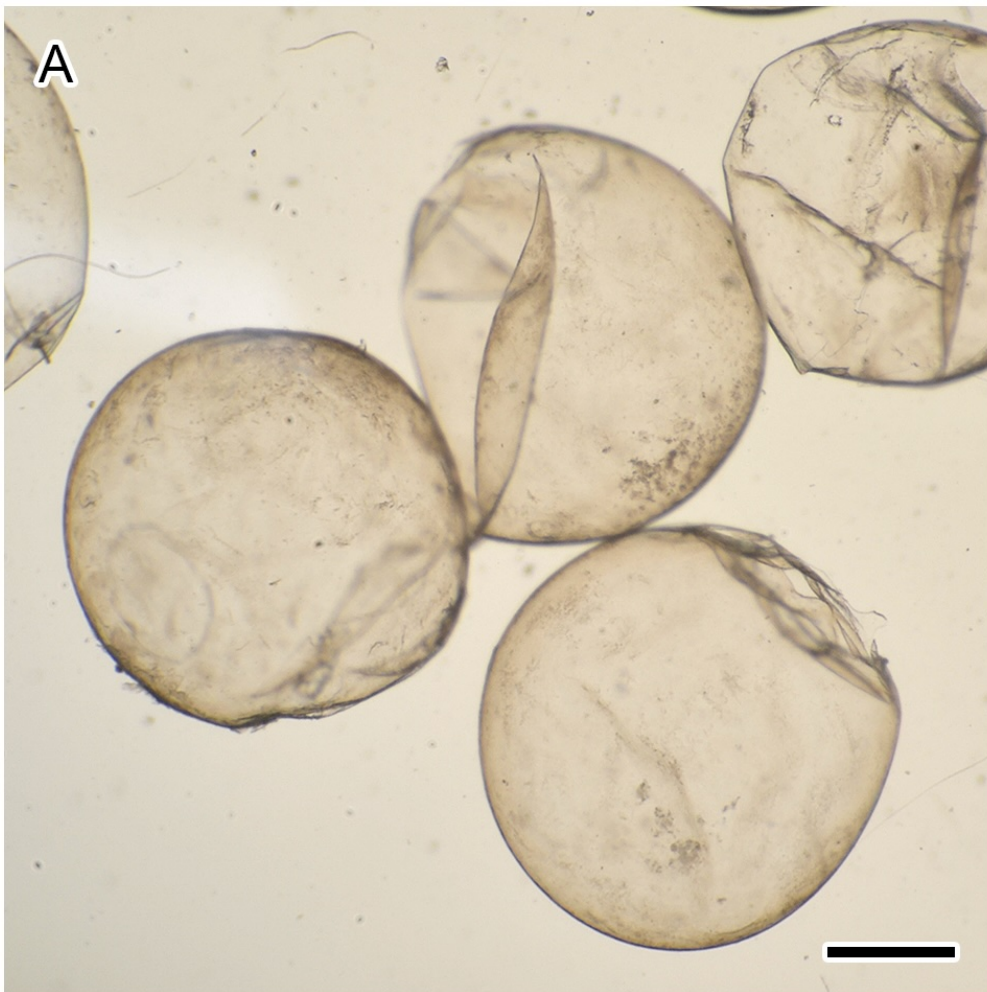
Supplementary Fig. 11. Allele frequency heatmap showing highly differentiated SNPs at the *ZPBA1* locus in a complete set of population samples of Pacific, Atlantic and Baltic herring. Purple boxes in the top row indicate missense positions.



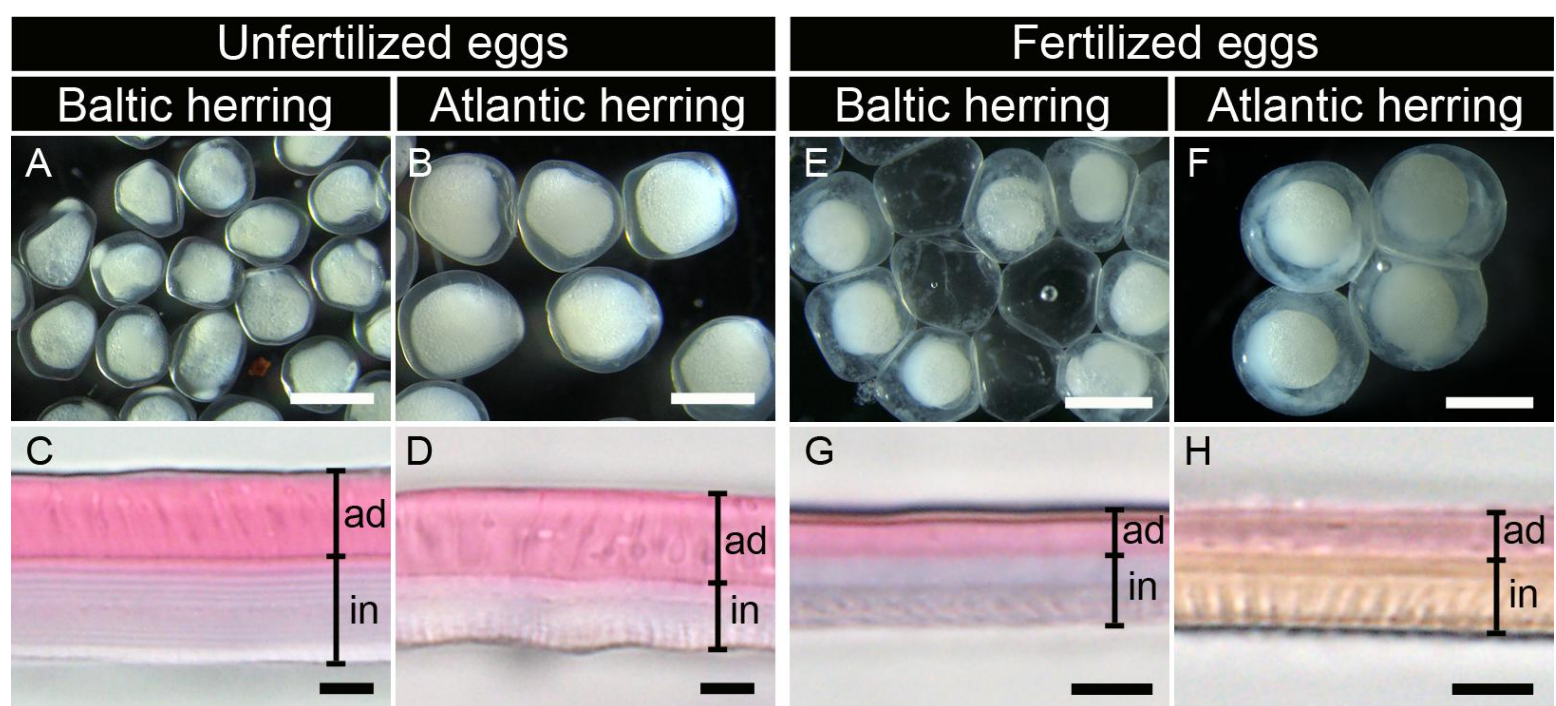
Supplementary Fig. S12. Individual haplotype neighbor-joining trees from the *FTG*, *ZPBA1*, and *HE1C* loci. (A) *FTG* (Chr 17:25.338 - 25.401 Mb); (B) *ZPBA1* (Chr 22:20.871 - 20.874 Mb); (C) *HE1C* (Chr 26:4.993 - 4.999 Mb). Branch lengths represent nucleotide divergence (i.e. average difference per base) across the included regions.



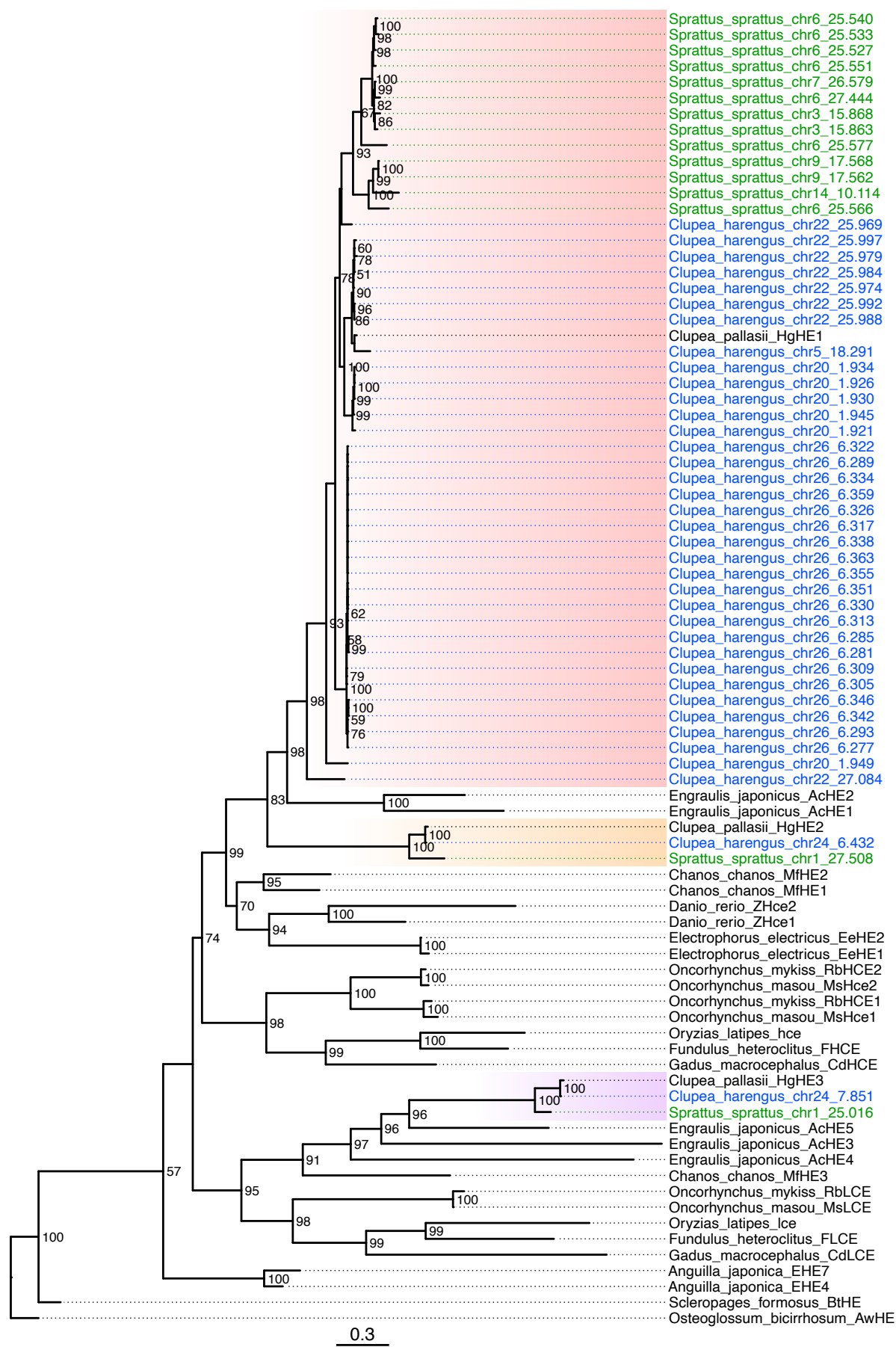
Supplementary Fig. 13. Clustering of individual haplotypes across the *FTG* locus. (A) Hierarchical clustering based on Hamming edit distance, where each edit corresponds to one nucleotide difference, between individual haplotypes across the *FTG* locus (chr 17: 25:34-25.40 Mb), showing that all Baltic haplotypes form a sub-clade inside the larger Pacific clade, without abnormal branch length. The Atlantic haplotypes found inside the Baltic clade stem from two coastal individuals caught at Askoy, Norway. (B) as (A), but based only on non-synonymous positions, showing that, on the amino acid level, there are Pacific haplotypes that are indistinguishable from those found in the Baltic. Also noteworthy is a single haplotype from Vancouver that is distinct from its peers, and instead cluster with the Subarctic/Baltic types.



Supplementary Fig. 14. Observation of isolated egg envelopes of Atlantic and Baltic herring in water droplets. Developing eggs of Atlantic and Baltic herring were crashed, and the egg envelopes were rinsed in PBS. The isolated egg envelopes of Atlantic (**A**) and Baltic herring (**B**) were placed in PBS droplets and observed under a binocular microscope. When water in the droplets was removed, the Atlantic egg envelopes collapsed (**C**), whereas the Baltic egg envelopes maintained a round shape (**D**), suggesting that Atlantic egg envelopes are softer than those of Baltic herring. Scale bar = 500 μ m.

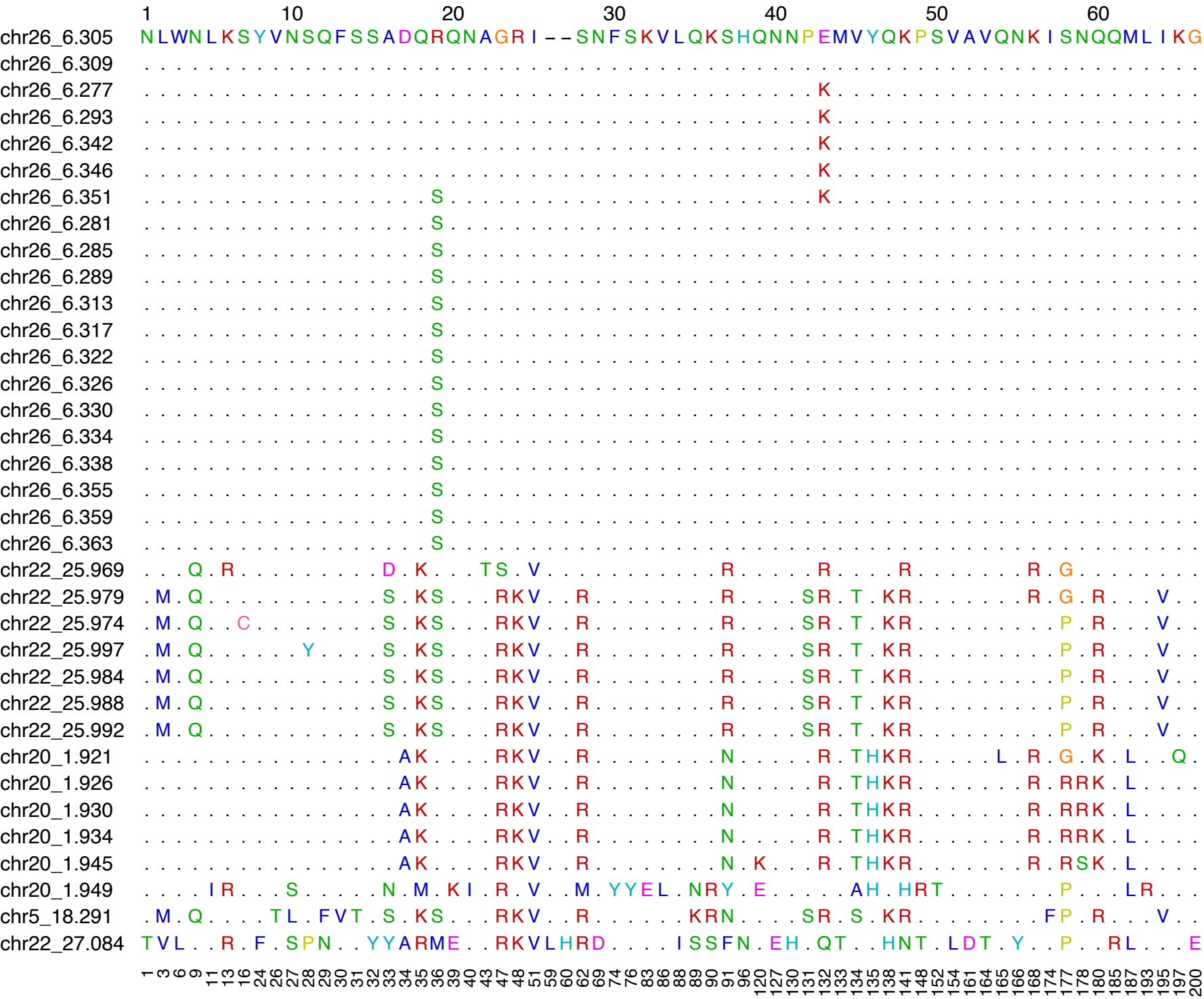


Supplementary Fig. 15. Observation of egg and egg envelope. Unfertilized (**A–D**) and fertilized (**E–H**) eggs of Baltic and Atlantic herring was compared. The 4% PFA fixed eggs were rehydrated and observed using a binocular microscope (**A, B, E, F**). HE-stained sections of the egg envelopes are shown in panels **C, D, G, H**. The lower side of the image corresponds to the cytoplasmic side. ad: adhesive layer, in: inner layer. Scale bars in (**A, B, E, F**): 1000 μm , and (**C, D, G, H**): 5 μm . The section of the egg envelopes was generated as follows. Fertilized and unfertilized eggs were fixed in 4% paraformaldehyde/PBS and stored in 70% ethanol at $-30\text{ }^{\circ}\text{C}$ until use. After stepwise replacement with PBS, the eggs were embedded in 5% agarose (Ultra-low Gelling Temperature Agarose, Sigma–Aldrich, MO, USA) prepared with 20% sucrose and then frozen. Frozen sections were cut at 14 μm thickness. After removal of agarose, sections were placed in Mayer’s Hematoxylin Solution (FUJIFILM Wako Chemical Co., Osaka, Japan) for 2 min, washed in running water for 15 min, and placed in 0.5% Eosin Y Solution (FUJIFILM Wako Chemical Co., Osaka, Japan) for 2 min. After dehydration with ethanol, they were placed in xylene for 15 min and sealed in LIMO mount (Pharma Co., Ltd., Tokyo, Japan).

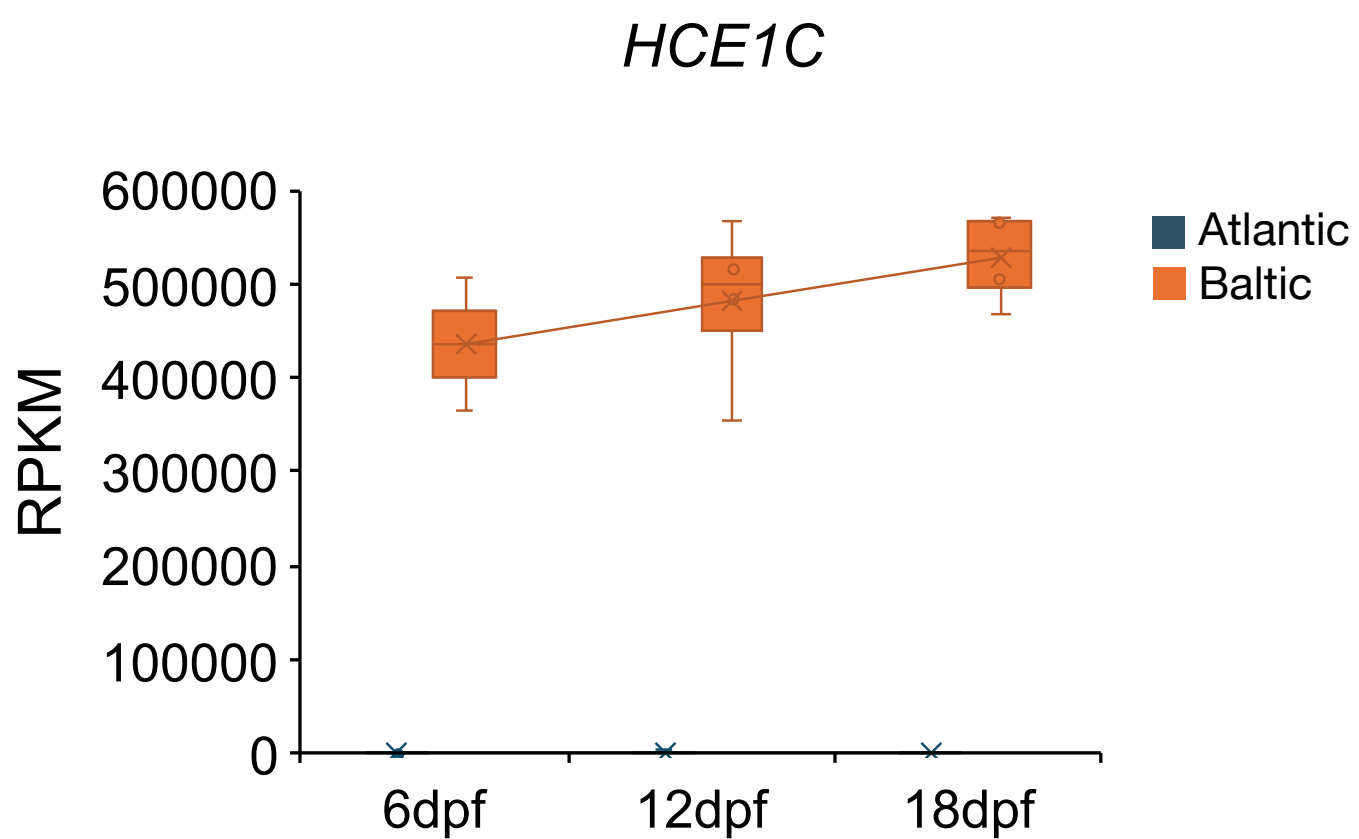


Supplementary Fig. 16. Maximum-likelihood phylogeny of fish hatching enzyme (HE) sequences from Atlantic herring and other fish.

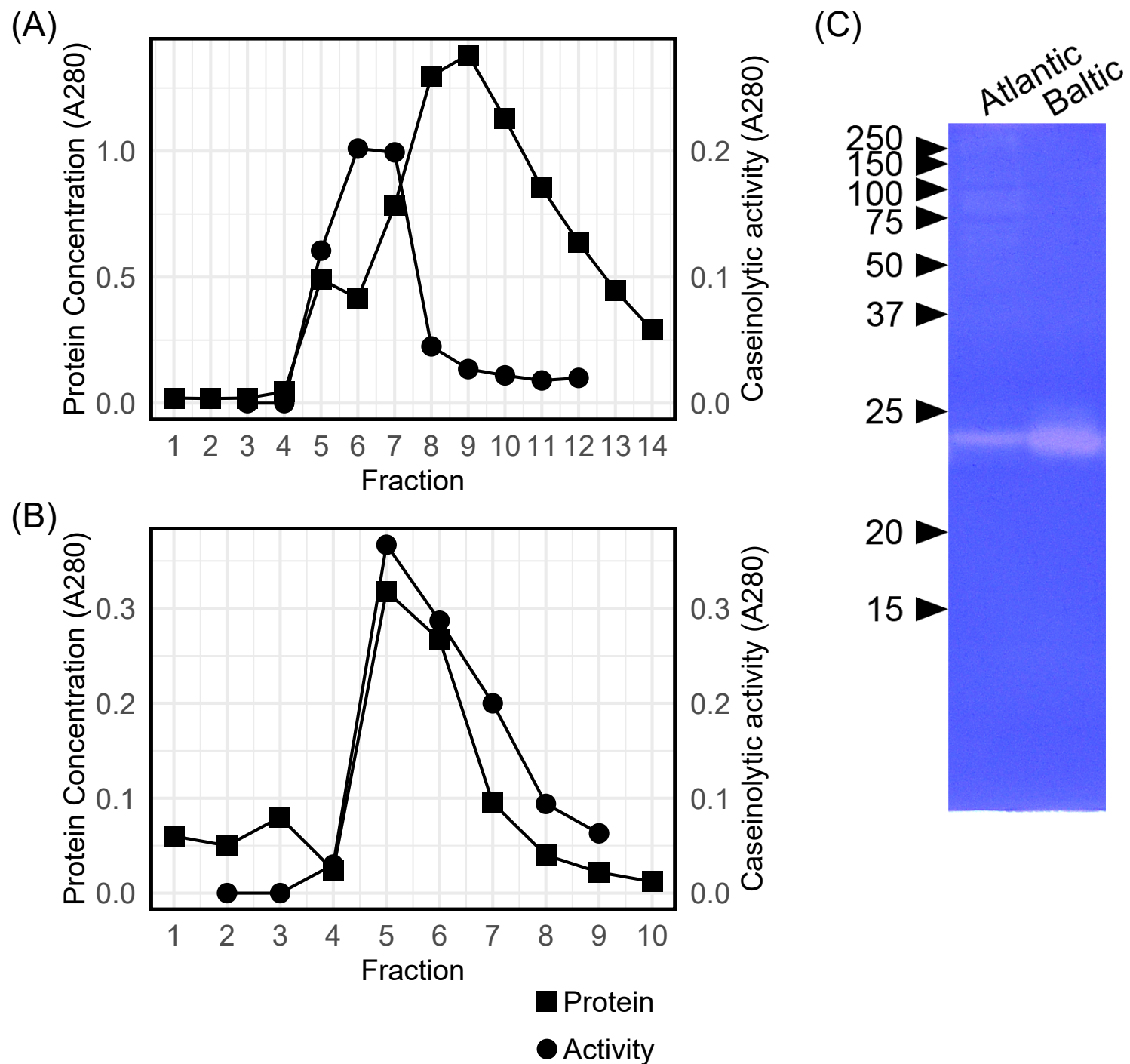
The tree was inferred with a codon substitution model using sequences corresponding to mature hatching enzyme proteins. Tip labels in blue and green denote Atlantic herring and European sprat sequences generated in this study, respectively. Tip labels for herring and sprat sequences follow the format Genus_species_chr#_position (Mb). Clades highlighted in red, orange and purple correspond to *HE1*, *HE2* and *HE3* genes in Clupeidae, respectively. Node values indicate bootstrap support from 2,000 replicates; values below 50% are not shown. The scale bar indicates branch length.



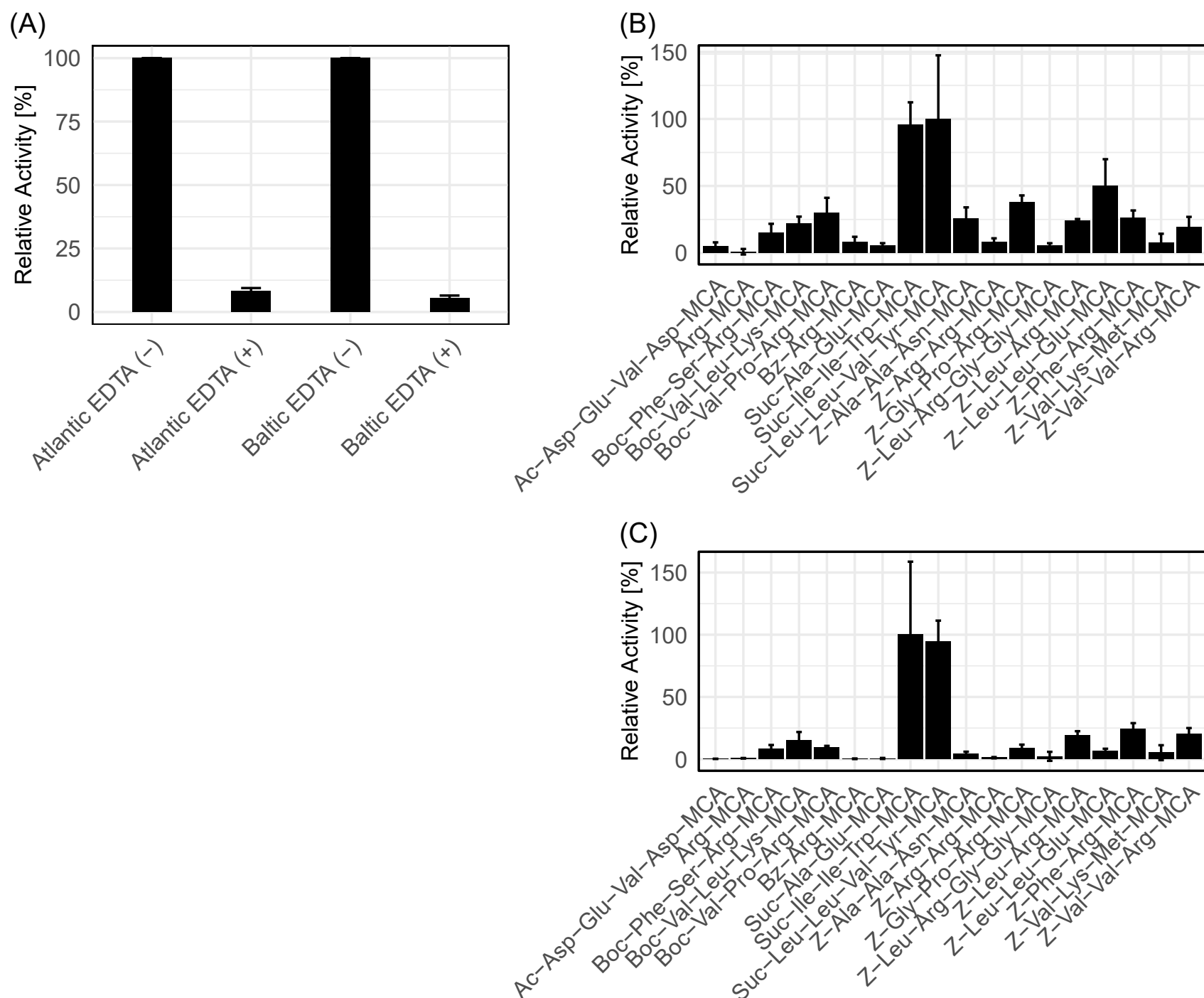
Supplementary Fig. 17. Amino acid polymorphism in the mature *HE1* protein of Atlantic herring. Amino acid sequence alignment showing only polymorphic sites within the mature *HE1* coding region identified in the Atlantic herring reference genome (Ch_v3.1). Sequence identifiers indicate genomic locations (chr#_position, Mb). A sequence from chromosome 26 is shown as the reference (top row); residues identical to the reference are denoted by dots, substitutions by letters, and gaps by dashes. Numbers above and below the alignment indicate variable-site indices and corresponding residue positions, respectively.



Supplementary Fig. 18. Expression of *HE1C* genes in developing embryos of Atlantic and Baltic herring. mRNA expression measured as RPKM (reads per kilobase million) in developing embryos 6, 12, and 18 days after fertilization. The minute expression indicated for Atlantic herring is most likely misalignment of reads from other HE loci.



Supplementary Fig. 19. Partial purification of hatching enzyme. One hundred mL of hatching liquid (culture medium collected after embryo hatching) from Atlantic or Baltic herring were dialyzed against 25 mM Tris-HCl (pH 7.5) containing 0.05% Brij 35 (Polyoxyethylene 23 Lauryl Ether), and applied to a Toyopearl SP-650M cation-exchange column (Tosoh Corp., Tokyo, Japan) equilibrated with the same buffer. The column was washed with the same buffer, and the adsorbed proteins were eluted in a single step using 25 mM Tris-HCl (pH 7.5) containing 1 M NaCl and 0.05% Brij 35. Panels (A) and (B) show the elution profiles of hatching enzymes from Atlantic and Baltic herring, respectively, on Toyopearl SP-650M column chromatography. (C) Casein zymography patterns of the partially purified hatching enzyme of Atlantic herring (left) and Baltic herring (right). The numbers on the left represent the size of the molecular markers. Each zymogram yielded a single band at 24 kDa, which is in agreement with values calculated from mature enzyme sequences of cDNAs (HE1-chr26, MW 22,657.0 and HE1-chr22, MW 22,911.3).



Supplementary Fig. 20. Inhibition of caseinolytic activity and substrate specificity of Atlantic and Baltic hatching enzymes. (A) Inhibition of caseinolytic activity of Atlantic and Baltic hatching enzymes. Activity is expressed as relative activity, with the activity in the absence of inhibitor (–) defined as 100%. The symbol (+) indicates the presence of EDTA. The partially purified enzyme was incubated with 10 mM EDTA solution for 10 min at 30 °C before testing caseinolytic activity as described above. (B, C) The substrate specificity of the partially purified hatching enzymes of Atlantic (B) and Baltic (C) herring examined using eighteen 4-methylcoumaryl-7-amide (MCA) peptides. The sequences of MCA peptides are shown on the X-axis. Activity toward each substrate is expressed as a percentage relative to the substrate showing the highest activity (set as 100%). MCA peptide cleavage activity was measured using a 50 μ L reaction mixture consisting of 50 mM Tris-HCl (pH 8.0), 100 μ M MCA peptide (Peptide Institute, Osaka, Japan), NaCl (0.25 M for Atlantic herring and 0 M for Baltic herring), and the enzyme. The mixture was incubated for 1 h at 30°C. The reaction was terminated by adding 100 μ L of 20% acetic acid. Fluorescence intensity of the solution was measured using a plate reader with excitation at 360 nm and emission at 460 nm. AMC standards (serial dilutions from 10 mM AMC, 1:1 to 1:1024) were used for calibration.

Supplementary Table 1: Region of genetic differentiation between Atlantic and Baltic spring-spawners. For each region, the section, if any, overlapping with the Ringkøbing vs Atlantic contrast is indicated, as are the genes highlighted in Figure 1. The region is defined as where SNPs with a delta allele frequency > 0.5 were detected in respective contrast. Regions with only one nucleotide position means that the region included only one SNP with delta allele frequency > 0.5.

Region (Atlantic vs Baltic)	Section shared with Ringkøbing	Gene of interest
chr1: 4,476,806		
chr1: 12,508,333 - 12,510,754	chr1: 12,510,753 - 12,510,754	
chr1: 13,537,685 - 13,540,866		
chr1: 17,792,091 - 17,988,019	chr1: 17,792,091 - 17,944,657	
chr1: 18,092,132 - 18,092,489		
chr1: 28,978,853		
chr2: 4,900,501		
chr2: 5,056,044 - 5,198,047	chr2: 5,058,121 - 5,193,008	<i>LRRC8C2</i>
chr4: 11,102,026 - 11,545,911		
chr8: 29,857,726 - 29,857,888		
chr9: 17,177,508 - 17,198,520		
chr9: 28,354,824 - 28,355,120	chr9: 28,354,824	
chr9: 28,503,413	chr9: 28,503,413	
chr10: 21,012,274	chr10: 21,012,274	
chr10: 21,304,351 - 21,676,835	chr10: 21,390,557 - 21,433,243	
chr10: 21,918,095 - 21,957,665		
chr10: 25,041,720 - 25,217,638	chr10: 25,087,106 - 25,212,675	<i>SEC16B</i>
chr10: 26,317,501		
chr12: 2,274,523 - 2,284,661	chr12: 2,277,917 - 2,277,939	<i>PRLRA</i>
chr12: 15,514,515 - 15,896,022	chr12: 15,667,590 - 15,896,022	<i>MYH</i>
chr12: 16,004,763 - 16,047,860	chr12: 16,004,763 - 16,033,289	<i>ENOPH1</i>
chr12: 25,698,579		
chr13: 5,752,250 - 5,811,490		
chr13: 19,664,381 - 19,667,152		
chr15: 10,871,077 - 10,912,023	chr15: 10,887,994 - 10,905,964	
chr16: 3,667,538 - 3,668,158		
chr16: 14,062,471 - 14,720,472	chr16: 14,161,023 - 14,647,635	<i>PLEKHA5</i>
chr16: 23,941,231 - 24,026,512		
chr16: 24,594,314		
chr16: 25,097,583		

Region (Atlantic vs Baltic)	Section shared with Ringkøbing	Gene of interest
chr17: 18,552,143 - 18,736,849	chr17: 18,573,486 - 18,734,586	
chr17: 25,290,121 - 25,438,174	chr17: 25,298,341 - 25,433,166	<i>FTG1-3</i>
chr18: 4,947,071 - 5,349,060	chr18: 5,119,712 - 5,276,576	<i>PGM5</i>
chr18: 24,356,394 - 24,378,542		
chr19: 6,359,267 - 6,398,947	chr19: 6,365,873 - 6,368,493	
chr21: 25,031,427	chr21: 25,031,427	
chr22: 20,855,359 - 20,879,304	chr22: 20,855,359 - 20,879,304	<i>ZPBA</i>
chr22: 22,912,073 - 22,927,743	chr22: 22,921,198 - 22,926,981	
chr24: 6,967,119		
chr24: 7,145,186		
chr26: 4,993,214 - 4,998,892	chr26: 4993214 - 4998892	<i>HE1C</i>
chr26: 8,588,254	chr26: 8,588,254	
unplaced_scaffold79: 112,517 - 112,603	unplaced_scaffold79: 112,546 - 112,603	
unplaced_scaffold101: 15,240		
unplaced_scaffold186: 68,818	unplaced_scaffold186: 68,818	
unplaced_scaffold319: 1,150		
unplaced_scaffold397: 19,597-32,070	unplaced_scaffold397: 19,597-32,070	
unplaced_scaffold728: 22,224		
unplaced_scaffold759: 1,906 - 22,677		
unplaced_scaffold787: 9,665 - 15,032	unplaced_scaffold787: 9,996 - 15,032	
unplaced_scaffold916: 20,729		

Supplementary Table 2. Nomenclature for *LRRC8C*, *ZPB1*, *FTG* and *HE* genes in the Atlantic herring genome. The existing nomenclature refers to the annotation of the Atlantic herring reference genome¹. Genes under strong selection as reported in this study are in bold.

Gene symbol			Location	
Proposed	Existing	Ensemble ID ¹	Chr	Region ¹
LRRC8C				
LRRC8C1	LRRC8C	ENSCHAG00020014387	10	10: 12,758,726-12,775,935
LRRC8C2	LRRC8C	ENSCHAG00020004192	2	2: 5,167,659-5,173,312
ZPB				
ZPBA1	zp2l2	ENSCHAG00020005751	22	22: 20,871,089-20,873,646
ZPBA2	NFYA	ENSCHAG00020007584	22	22: 20,987,288-20,989,935
ZPBB	zp2l2	ENSCHAG00020031478	12	12: 15,327,382-15,330,990
FTG & F13A				
FTG1	F13A1	ENSCHAG00020011400	17	17: 25,337,999-25,346,141
FTG2	F13A1	ENSCHAG00020011411	17	17: 25,359,224-25,369,735
FTG3	F13A1	ENSCHAG00020011471	17	17: 25,388,162-25,401,039
FTG4	-	ENSCHAG00020010975	15	15: 21,706,688-21,711,513
FTG5	-	ENSCHAG00020010977	15	15: 21,734,309-21,745,130
F13A1	f13a1b	ENSCHAG00020005056	18	18: 13,116,789-13,129,741
HE				
HE1A	c6ast1	ENSCHAG00020007338	20	20: 1,820,389-1,824,457
HE1A	-	ENSCHAG00020007375	20	20: 1,828,364-1,830,987
HE1A	-	ENSCHAG00020007379	20	20: 1,832,649-1,834,596
HE1A	-	ENSCHAG00020016857	20	20: 1,838,575-1,841,522
HE1A	-	ENSCHAG00020017181	20	20: 1,843,215-1,845,798
HE1A	-	ENSCHAG00020016360	20	20: 1,847,134-1,849,408
HE1A	-	ENSCHAG00020016584	20	20: 1,860,315-1,863,002
HE1A	SOST	ENSCHAG00020016921	20	20: 1,872,563-1,874,837
HE1A	-	ENSCHAG00020017229	20	20: 1,875,798-1,877,829
HE1B	npsn	ENSCHAG00020002248	22	22: 23,437,653-23,441,458
HE1B	-	ENSCHAG00020002285	22	22: 23,443,301-23,445,308
HE1B	-	ENSCHAG00020002287	22	22: 23,447,669-23,449,679
HE1B	-	ENSCHAG00020002290	22	22: 23,452,070-23,455,343
HE1B	-	ENSCHAG00020002421	22	22: 23,457,572-23,459,325
HE1B	-	ENSCHAG00020002428	22	22: 23,461,721-23,463,728
HE1B	-	ENSCHAG00020002447	22	22: 23,466,201-23,468,533
HE1B	-	ENSCHAG00020002768	22	22: 23,470,856-23,472,867
HE1C	-	ENSCHAG00020030368	26	26: 4,974,200-4,976,139
HE1C	-	ENSCHAG00020030549	26	26: 4,978,487-4,980,243
HE1C	-	ENSCHAG00020030553	26	26: 4,982,567-4,984,386
HE1D	-	ENSCHAG00020014953	5	5: 18,034,408-18,037,713
HE1E	-	ENSCHAG00020002224	22	22: 24,602,480-24,604,878
HE2	-	ENSCHAG00020012460	24	24: 6,543,146-6,546,666
HE3	-	ENSCHAG00020009635	24	24: 8,049,184-8,052,630

¹Ensembl genome assembly: Ch_v2.0.2v2 (GCA_900700415.2); Annotation: Ensembl release 115, 2025-07-13

Supplementary Table 3. Species and genomic resources used for *LRRC8C* analyses.

Species (English name)	Family	Dataset tag	Genome assembly Ensembl (NCBI)	Genome annotation Ensembl/NCBI (version)	Annotation source	Nr of protein coding genes n (BUSCO)
<i>Denticeps clupeioides</i> (Denticle herring)	Denticipitidae	dclu	FDenClu1.1 (GCA_900700375.1)	Ensembl: fDenClu1.1 (109)	https://ftp.ensembl.org/pub/release-109	24,142 (C:93.4%[S:86.3%,D:7.1%],F:1.8%,M:4.8%,n:3640)
<i>Coilia nasus</i> (Japanese grenadier anchovy)	Engraulidae	cnas	(GCA_027475355.1; ASM2747535v1)	n/a*	<i>de novo</i> using <i>C. harengus</i> models	21,432 (85.3%[S:82.3%,D:3.0%],F:2.4%,M:12.3%,n:3640)
<i>Clupea harengus</i> (Atlantic herring)	Clupeidae	char	Ch_v2.0.2	Ensembl: Ch_v2.0.2 (109)	https://ftp.ensembl.org/pub/release-109	24,095 (C:88.6%[S:86.0%,D:2.6%],F:2.1%,M:9.3%,n:3640)
<i>Sprattus sprattus</i> (European sprat)	Clupeidae	sspr	Sprat_DeDup_v2_HiC (GCA_041430245.1; ASM4143024v1)	n/a*	<i>de novo</i> using <i>C. harengus</i> models	22,223 (C:83.5%[S:80.6%,D:2.9%],F:2.6%,M:13.9%,n:3640)
<i>Sardina pilchardus</i> (European pilchard)	Alosidae	spil	(GCA_900499035.1)	n/a*	<i>de novo</i> using <i>A. sapidissima</i> models	23,494 (C:77.7%[S:75.2%,D:2.5%],F:6.4%,M:15.9%,n:3640)

<i>Alosa alosa</i> *** (American shad)	Alosidae	aalo	AALO_Geno_1.1 (GCF_017589495.1)	NCBI RefSeq: NCBI Alosa alosa Annotation Release (100)	https://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_other/Alosa_alosa/annotation_releases/100/	23,465 (C:94.7%[S:92.4%,D:2.3%],F:1.1%,M:4.2%,n:3640)
<i>Alosa sapidissima</i> (American shad)	Alosidae	asap	fAloSap1.pri (GCF_018492685.1)	NCBI RefSeq: fAloSap1 (100)	https://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate_other/Alosa_sapidissima/annotation_releases/100/	25,717 (C:98.6%[S:96.3%,D:2.3%],F:0.5%,M:0.9%,n:3640)
<i>Limnothrissa miodon</i> (Lake Tanganyika sardine)	Dorosomatidae	lmio	(GCA_017657215.1; ASM1765721v1)	n/a*	<i>de novo</i> using <i>A. sapidissima</i> models	20,951 (C:83.1%[S:79.6%,D:3.5%],F:2.9%,M:14.0%,n:3640)
<i>Tenualosa ilisha</i> (Hilsa shad)	Dorosomatidae	tili	(GCA_015244755.2)	n/a*	<i>de novo</i> using <i>A. sapidissima</i> models	23,529 (C:97.6%[S:95.5%,D:2.1%],F:0.8%,M:1.6%,n:3640)

*No public genome annotation available in Ensembl or NCBI at the time of analysis or submission.

**BUSCO completeness scores: C:%Complete[S:%Complete and single-copy,D:%Complete and duplicated],F:%Fragmented,M:%Missing,n:Total BUSCO groups searched.

***Species used in orthogroup clustering but excluded from downstream phylogenetic analyses.

Supplementary Table 4. Overview of LRRC8C1 and LRRC8C2 peptides monitored in sperm and tissue samples, whether presence could be verified and the corresponding limit of detection determination.

Tissue/Sperm	Protein	Peptide	Expression	Limit of Detection
Sperm	LRRC8C1	LYIYNDGTK	Not verified	0.005 fmol/1 µg
Sperm	LRRC8C1	YLDLSYNDIR	Not verified	0.05 fmol/1 µg
Sperm	LRRC8C2	VHVEEGNLLYK	Verified	n.a.
Sperm	LRRC8C2	LGNNLLSGLSPK	Verified	n.a.
Sperm	LRRC8C2	LYFSHNK	Verified	n.a.
Sperm	LRRC8C2	PVAVASVLDSK	Verified	n.a.
Liver	LRRC8C2	VHVEEGNLLYK	Not verified	0.005 fmol/1 µg
Liver	LRRC8C2	LGNNLLSGLSPK	Not verified	0.1 fmol/1 µg
Liver	LRRC8C2	LYFSHNK	Not verified	1 fmol/1 µg
Liver	LRRC8C2	PVAVASVLDSK	Not verified	0.005 fmol/1 µg
Muscle	LRRC8C2	VHVEEGNLLYK	Not verified	0.005 fmol/1 µg
Muscle	LRRC8C2	LGNNLLSGLSPK	Not verified	0.005 fmol/1 µg
Muscle	LRRC8C2	LYFSHNK	Not verified	0.05 fmol/1 µg
Muscle	LRRC8C2	PVAVASVLDSK	Not verified	0.005 fmol/1 µg
Brain	LRRC8C2	VHVEEGNLLYK	Not verified	0.005 fmol/1 µg
Brain	LRRC8C2	LGNNLLSGLSPK	Not verified	0.1 fmol/1 µg
Brain	LRRC8C2	LYFSHNK	Not verified	0.1 fmol/1 µg
Brain	LRRC8C2	PVAVASVLDSK	Not verified	0.005 fmol/1 µg
Heart	LRRC8C2	VHVEEGNLLYK	Not verified	0.005 fmol/1 µg
Heart	LRRC8C2	LGNNLLSGLSPK	Not verified	0.1 fmol/1 µg
Heart	LRRC8C2	LYFSHNK	Not verified	1 fmol/1 µg
Heart	LRRC8C2	PVAVASVLDSK	Not verified	0.005 fmol/1 µg

Supplementary Table 5. The selected proteins from the Atlantic herring sperm depicted in the ranking plot (Fig. 2e). iBAQ values and other information are taken from the MaxQuant output (Dataset S1).

Rank by iBAQ	Name in the ranking plot	Protein IDs	Mol. weight [kDa]	Sequence length	iBAQ peptides	iBAQ
14	Na/K-ATPase α	A0A6P8GSB7	113.3	1027	49	4.84E+09
34	PKA	A0A6P3VVV0; A0A6P8F597; A0A6P8ESG0; A0A6P8F2V2	40.551	351	18	3.62E+09
56	PKA RII β	A0A6P8GGN7; A0A6P8GL57	43.53	385	23	2.98E+09
104	HCN channel 1	A0A6P3WFL3; A0A6P8GIJ0	67.455	593	22	1.8E+09
106	Ca ²⁺ ATPase	A0A6P3VS68; A0A6P8GNB1; A0A6P8GJ19	136.97	1245	45	1.79E+09
207	HCN channel 2	A0A8M1KM12; A0A8M1KNY3	52.562	457	17	8.42E+08
269	SPACA9	A0A6P3VJL4	26.444	236	13	5.74E+08
282	sNHE (SLC9C1)	A0A6P8F5G0	133.06	1177	49	5.38E+08
407	Adenylyl cyclase 10 (sAC)	A0A8M1KSD9	193.16	1704	94	3.26E+08
440	NCKX2 (SLC24A2)	A0A6P8GMS5; A0A6P8GNY6; A0A6P8GY85; A0A6P8H080	68.904	628	16	2.96E+08
679	HCN channel 3	A0A8M1KW70	33.467	291	15	1.4E+08
732	CatSper β	A0A6P8GJ87	15.807	142	10	1.18E+08
777	CatSper δ	A0A8M1KQ62; A0A6P8FLB3; A0A8M1KML3	82.206	727	31	1.05E+08
810	LRRC8C2 BalticN*	LRRC8C2_BalticN	91.072	793	43	96,158,000
937	CatSper ϵ	A0A6P8EG77	70.908	614	34	69,319,000
991	Slo1-like channel	A0A6P8FLH9	68.599	614	28	62,550,000
2218	PKA RI α	A0A8M1KHN9	42.934	379	20	6,989,700
2234	DC-STAMP2	A0A6P3VQE2	89.616	795	39	6,837,600
2556	SLC2 (GLUT)	A0A6P8GFD1	63.956	596	16	4,232,700
3074	ACAP9	A0A6P8H4B3	508.74	4475	268	1,793,900
3111	Izumo-1	A0A8M1KNY4	38.676	342	14	1,683,000
3163	SPACA4	A0A6P3W1J8	13.84	129	6	1,5303,00
3223	TMEM81	A0A6P3W5N0	35.36	321	18	1,393,800

3447	LRRC8C2**	A0A6P3VMJ8; LRRC8C2_Intermediate; LRRC8C2_BalticS	91.059	793	42	956,650
3821	NKCC1 (SLC12A2)	A0A6P8FL89	122.08	1109	44	470,320
3988	AKAP1b	A0A6P8FN49	101.03	936	33	325,870
4336	AKAP8	A0A6P8F493	85.306	821	41	103,760

* Baltic North allele product identified based on two diagnostic (proteotypic) peptides.

** Protein group including Atlantic, Intermediate or Baltic south variants are indistinguishable.

Supplementary Table 6. Pore residues in mouse LRRC8C and herring LRRC8C1 and LRRC8C2 corresponding to residues K98, D100, R103, and H104 in human LRRC8A.

Species	Residue(s)			
<i>H. sapiens</i> LRRC8A	K98	D100	R103	H104
<i>M. musculus</i> LRRC8C	K100	D102	L105	Q106
<i>C. harengus</i> C1	K102	N104	W107	Q108
<i>C. harengus</i> C2	T99	N101	L104	Q105

Supplementary Table 7. Sequences from other vertebrate species homologous to *FTG* and *F13A* genes in Atlantic herring, used in the phylogenetic analysis.

Scientific name	Common name	Gene	Accession number/ Gene ID
<i>Astyanax mexicanus</i>	Mexican tetra	<i>FTG</i>	XP_022529121.2
<i>Astyanax mexicanus</i>	Mexican tetra	<i>F13A</i>	XP_007237893.3
<i>Danio rerio</i>	Zebrafish	<i>FTG</i>	XP_686649.5
<i>Danio rerio</i>	Zebrafish	<i>FTG</i>	NP_001070179.2
<i>Danio rerio</i>	Zebrafish	<i>F13A</i>	NP_001070622.1
<i>Electrophorus electricus</i>	Electric eel	<i>FTG</i>	XP_026869086.2
<i>Electrophorus electricus</i>	Electric eel	<i>F13A</i>	XP_026882133.2
<i>Homo sapiens</i>	Human	<i>F13A</i>	NP_000120.2
<i>Mus musculus</i>	House mouse	<i>F13A</i>	NP_001159863.1
<i>Oreochromis niloticus</i>	Nile tilapia	<i>FTG</i>	XP_003452423.1
<i>Oreochromis niloticus</i>	Nile tilapia	<i>F13A</i>	XP_019208965.1
<i>Oryctolagus cuniculus</i>	Rabbit	<i>F13A</i>	XP_051711102.1
<i>Oryzias latipes</i>	Japanese medaka	<i>FTG</i>	XP_023806131.1
<i>Oryzias latipes</i>	Japanese medaka	<i>F13A</i>	XP_011485797.1
<i>Alosa sapidissima</i>	American shad	<i>FTG</i>	XP_041923962.1
<i>Alosa sapidissima</i>	American shad	<i>F13A</i>	XP_041925040.1
<i>Denticeps clupeoides</i>	Denticle herring	<i>FTG</i>	ENSDCDG00000005193.1
<i>Denticeps clupeoides</i>	Denticle herring	<i>F13A</i>	ENSDCDG000000021470.1
<i>Limnothrissa miodon</i>	Lake Tanganyika sardine	<i>FTG</i>	N/A*
<i>Limnothrissa miodon</i>	Lake Tanganyika sardine	<i>F13A</i>	N/A*
<i>Sardina pilchardus</i>	European pilchard	<i>FTG</i>	XP_062377426.1
<i>Sardina pilchardus</i>	European pilchard	<i>FTG</i>	XP_062377575.1
<i>Sardina pilchardus</i>	European pilchard	<i>FX13A</i>	XP_062372726.1
<i>Sprattus sprattus</i>	European sprat	<i>FTG</i>	N/A*
<i>Sprattus sprattus</i>	European sprat	<i>F13A</i>	N/A*
<i>Tenualosa ilisha</i>	Hilsa shad	<i>FTG</i>	N/A*
<i>Tenualosa ilisha</i>	Hilsa shad	<i>F13A</i>	N/A*

*N/A: Sequences annotated *de novo* in this study (see Supplementary Table 3).

Supplementary Table 8. Average d_N , d_S and d_N/d_S ratios for different branch groups based on Fig. 3A*.

Branch group	N*	d_N	d_S	d_N/d_S
<i>FTG</i>				
Atlantic herring	8	0.024	0.012	2.417
Clupeiformes	22	0.065	0.142	1.331
Clupeiformes excluding herring	14	0.089	0.215	0.710
Other teleost	10	0.137	0.501	0.375
<i>F13A</i>				
Atlantic herring	1	0.006	0.044	0.131
Clupeiformes	13	0.033	0.183	0.200
Other teleost	8	0.061	0.395	0.172
Mammal	3	0.032	0.333	0.112

*Branches marked with a minus sign in Fig. 3A were excluded from the calculations.

Supplementary Table 9. The selected proteins from the Atlantic herring oocyte depicted in ranking plot (Fig. 3E). iBAQ values and other information are taken from the MaxQuant output table (Dataset S1).

Rank by iBAQ	Name in the ranking plot (Figure 3e)	Protein IDs	Mol. weight [kDa]	Sequence length	iBAQ peptides	iBAQ
2	ZPCA ¹	A0A8M1K9Z8	44.296	408	15	11,908,000,000
7	Nucleoplasmin-2b	A0A6P3VE06	21.054	191	7	7,848,600,000
8	ZPBA2 ²	A0A6P8EH72	58.756	534	24	6,916,600,000
22	ZPBA1	A0A6P3VLM4; A0A8M1K577	55.158	501	24	4,336,500,000
448	FTG1	A0A8M1KU37	71.965	639	33	259,600,000
480	FTG3	A0A6P8GKT3	81.997	727	40	233,470,000
495	FTG2	A0A6P8GP96	59.704	528	28	229,040,000
844	Tetraspanin (CD9b)	A0A6P3VTJ1; A0A6P3VTY5	24.904	229	7	98,508,000
1363	ZPBB	A0A6P8G2U1	45.601	413	16	40,580,000
1373	FTG5	A0A6P8GCR7	79.64	712	36	40,208,000
2236	Nucleoplasmin-3	A0A8M1KFL0	17.253	158	5	13,250,000
3181	Tetraspanin (CD9a)	A0A6P3VKJ1; A0A6P3VK17	25.054	227	8	4,957,600
3392	Protein Bouncer	A0A6P8ESF9	16.979	155	6	4,106,000
3928	Anoctamin Cl ⁻ channel	A0A6P3WCF7; A0A6P8GBX7; A0A6P3WCB0; A0A6P8GAJ6	75.536	655	22	2,471,600
4771	FTG4	A0A6P8GK41	25.776	225	13	1,060,300

^{1,2}The corresponding genes are annotated as *ZP3* and *ZP4-like*, respectively, in the Atlantic herring assembly Ch_v2.0.2v2 (GCA_900700415.2); Annotation: Ensembl release 115, 2025-07-13.

Supplementary Table 10. Results of *in vitro* fertilization of Atlantic and Baltic herring at different salinities

Source	Female id.	Salinity (‰)	Replicates	Fertilization rate (%)
Atlantic	1	6	4	20
Atlantic	1	16	3	95
Atlantic	1	35	4	95
Atlantic	9	6	20	30
Atlantic	9	16	3	40
Atlantic	9	35	20	2**
Atlantic	10	6	4	20
Atlantic	10	16	3	85
Atlantic	10	35	4	80
Atlantic	15	16	5	95
Baltic	1	6	1	95
Baltic	3	6	12	85
Baltic	3	16	2	45
Baltic	3	35	2	3
Baltic	5	6	2	10
Baltic	5	16	2	65
Baltic	5	35	2	1
Baltic	6	6	2	20
Baltic	6	16	3	70
Baltic	6	35	2	90*
Baltic	9	16	1	95
Baltic	10	16	1	70
Baltic	11	16	1	95
Baltic	12	16	1	95
Baltic	13	16	1	95
Baltic	14	16	1	70
Baltic	15	6	2	5
Baltic	15	16	3	73
Baltic	15	35	2	5

* Most embryos with developmental error

** Unexpected low fertilization rate

Supplementary Table 11. Protein density of egg envelope determined by egg diameter, thickness of inner layer and protein amount per egg.

	Unfertilized		<i>P</i> -value (Welch <i>t</i>)	Fertilized		<i>P</i> -value (Welch <i>t</i>)
	Baltic	Atlantic		Baltic	Atlantic	
Egg diameter (μm) ^a	927 ± 7.65	1354 ± 9.98	1.9 × 10 ⁻⁸	1292 ± 11.7	1574 ± 6.26	9.6 × 10 ⁻⁷
Thickness (μm) ^b	10.48 ± 0.32	7.33 ± 0.18	3.2 × 10 ⁻⁵	4.29 ± 0.09	3.93 ± 0.15	8.1 × 10 ⁻³
Protein per egg (μg) ^c				31.6 ± 0.81	20.8 ± 1.26	4.5 × 10 ⁻⁵
Volume (×10 ⁻² mm ³) ^d	2.77 ± 0.010	4.17 ± 0.12	7.9 × 10 ⁻⁵	2.23 ± 0.061	3.04 ± 0.12	1.7 × 10 ⁻⁴
Density (μg/mm ³) ^e	1143 ± 48.8	498 ± 33.4	6.0 × 10 ⁻⁶	1414 ± 52.8	685 ± 49.2	4.0 × 10 ⁻⁶

^a The diameter of the inner layer, not including the adhesive layer, was measured, n=3

^b The thickness of egg envelopes was measured as the inner layer of egg envelopes in the sections of 4% PFA fixed embryos. n=3

^c See methods for how the amount of protein per egg envelope was measured.

^d Volume of egg envelope protein (*V*) were estimated as follows.

$$V = \frac{4}{3} \pi r^3 - \frac{4}{3} \pi (r - t)^3$$

where *r* represents the egg radius, and *t* represents the egg envelope thickness

^e The density of egg envelope proteins is expressed as protein per egg / egg envelope volume.

Supplementary Table 12. Hatching enzyme sequences from teleost species used as queries for gene annotation and for phylogenetic reconstruction.

Common name	Scientific name	Gene name	Accession No.	Source
Pacific herring	<i>Clupea pallasii</i>	HgHE1	AB433584	mRNA
		HgHE2	AB433585	mRNA
		HgHE3	AB433586	mRNA
Japanese anchovy	<i>Engraulis japonicus</i>	AcHE1	AB433587	mRNA
		AcHE2	AB433588	mRNA
		AcHE3	AB433589	mRNA
		AcHE4	AB433590	mRNA
		AcHE5	AB433591	mRNA
milkfish	<i>Chanos chanos</i>	MfHE1	AB480009	mRNA
		MfHE2	AB480010	mRNA
		MfHE3	AB480011	mRNA
zebrafish	<i>Danio rerio</i>	ZHce1	AB175621	mRNA
		ZHce2	AB175620	mRNA
electric eel	<i>Electrophorus electricus</i>	EeHE1	AB480019	genomic DNA
		EeHE2	AB480020	genomic DNA
masu salmon	<i>Oncorhynchus masou</i>	MsHce1	AB175619	mRNA
		MsHce2	AB175618	mRNA
		MsLCE	AB480021	genomic DNA
rainbow trout	<i>Oncorhynchus mykiss</i>	RbHCE1	AB480022	mRNA
		RbHCE2	AB480023	mRNA
		RbLCE	AB480024	mRNA
Pacific cod	<i>Gadus macrocephalus</i>	CdHCE	AB480029	mRNA
		CdLCE	AB480030	mRNA
killifish	<i>Fundulus heteroclitus</i>	FHCE	AB210813	mRNA
		FLCE	AB210814	mRNA
medaka	<i>Oryzias latipes</i>	MHCE	M96170	mRNA
		MLCE	M96169	mRNA
Japanese eel	<i>Anguilla japonica</i>	EHE4	AB071423	mRNA
		EHE7	AB071425	mRNA
Asian bonytongue	<i>Scleropages formosus</i>	BtHE	AB480003	genomic DNA
arowana	<i>Osteoglossum bicirrhosum</i>	AwHE	AB276000	genomic DNA

All entries in this table are sourced from Kawaguchi et al. (2010).

Supplementary Table 13. Analysis of copy number variation at the *HEIC* locus on chromosome 26 in different herring populations based on short read, pooled whole genome resequencing.

ID	Sample_ID	Location	Group	CNV	Salinity (‰)
1	Kalix_Baltic_Spring	Kalix	Baltic	23	2.4
2	Riga_Baltic_Spring	Gulf of Riga	Baltic	26	4.1
3	Riga_Baltic_Spring	Gulf of Riga	Baltic	27	4.1
4	Riga_Baltic_Autumn	Gulf of Riga	Baltic	24	5.2
5	Riga_Baltic_Autumn	Gulf of Riga	Baltic	20	5.2
6	Gävle_Baltic_Summer	Gävle	Baltic	22	3.7
7	Vaxholm_Baltic_Spring	Vaxholm	Baltic	22	5.8
8	Hästkär_Baltic_Spring	Hästkär	Baltic	24	4.0
9	Hudiksvall_Baltic_Spring	Hudiksvall	Baltic	21	4.5
10	Gävle_Baltic_Spring	Gävle	Baltic	20	5.0
11	Gävle_Baltic_Autumn	Gävle	Baltic	19	6.0
12	Gamleby_Baltic_Spring	Gamleby	Baltic	21	6.5
13	Kalmar_Baltic_Spring	Kalmar	Baltic	24	6.5
14	Karlskrona_Baltic_Spring	Karlskrona	Baltic	23	6.7
15	Rügen_Baltic_Spring	Rügen	Baltic	21	5.8
16	Rügen_Baltic_Spring	Rügen	Baltic	31	4.0
17	CentralBaltic_Baltic_Spring	Central Baltic	Baltic	21	7.2
18	Germany_Baltic	Ariadnegrund	Baltic	25	5.0
18	Germany_Baltic	Ariadnegrund	Baltic	25	5.0
19	BornholmBasin_Baltic_Autumn	Bornholm Basin	Baltic	17	7.7
20	Schlei_Baltic_Spring	Schlei	Transition	25	9.0
21	Schlei_Baltic_Autumn	Schlei	Transition	27	9.0
22	Fehmarn_Baltic_Autumn	Fehmarn	Transition	16	12.0
23	RingkobingFjord_NorthSea_Spring	Ringköbing Fjord	Atlantic	24	12.0

24	Landvik_Atlantic_Spring	Landvik	Transition	16	15.0
25	Norway_Baltic_Spring	Landvik	Transition	17	15.0
26	Träslövsläge_Baltic_Spring	Träslövsläge	Transition	22	10.0
27	Kattegat_Atlantic_Spring	Kattegat, Björköfjorden	Transition	16	18.4
28	Hamburgsund_Atlantic_Spring	Hamburgsund	Transition	7	25.4
29	KattegatNorth_Atlantic_Spring	Kattegat North	Transition	9	25.4
30	Skagerrak_Atlantic_Spring	Skagerrak, Brofjorden	Transition	8	25.0
31	Lindås_Atlantic_Spring	Lindås	Atlantic	5	28.0
32	Lusterfjorden_Atlantic_Spring	Lusterfjorden	Atlantic	2	32.0
33	Clyde_Atlantic_Spring	Ballantrae. Clyde	Atlantic	1	33.0
34	IsleOfMan_IrishSea_Autumn	Douglas Bank. Isle of Man	Atlantic	1	33.0
35	TeelinBay_Atlantic_Winter	Teelin Bay	Atlantic	0	34.0
36	Downs_EnglishChannel_Winter	Downs	Atlantic	1	35.0
37	Greenland_Atlantic_Spring	Greenland	Atlantic	0	35.0
38	Atlantic_Spring	Fortune Bay	Atlantic	0	35.0
39	Atlantic_Spring	Inner Baie Des Chaleurs	Atlantic	0	35.0
40	Atlantic_Spring	Northumberland Strait	Atlantic	0	35.0
41	NSSH_Atlantic_Spring	Norway	Atlantic	0	35.0
42	CapeWrath_Atlantic_Spring	Isle of Skye	Atlantic	0	35.0
43	Gloppen_Atlantic_Spring	Gloppefjorden	Atlantic	2	35.0
44	Norway_Atlantic_Atlantic_Spring	Norway	Atlantic	0	35.0
45	Iceland_Atlantic_Spring	Iceland. Höfn	Atlantic	1	35.0
46	Hebrides_Atlantic_Mixed	West of Hebrides	Atlantic	1	35.0
47	CelticSea_Atlantic_AutumnWinter	Celtic Sea	Atlantic	2	35.0
48	DalBoB_Atlantic_Autumn	Bonavista Bay	Atlantic	0	35.0
49	DalGeB_Atlantic_Autumn	German Banks	Atlantic	0	35.0
50	DalNsF_Atlantic_Autumn	Northumberland Strait	Atlantic	0	35.0
51	Orkney_NorthSea_Autumn	Orkney	Atlantic	1	35.0
52	CapeWrath_Atlantic_Autumn	Cape Wrath	Atlantic	1	35.0

53	NorthSea_Atlantic_Autumn	North Sea	Atlantic	0	35.0
54	Pacific_Pacific_Spring	Vancouver, Strait of Georgia	Pacific	8	25.0
55	Japan_Sea of Japan	Sea of Japan, Sakhalin	Pacific	18	15.0
56	LandvikS32_Norway_Baltic_Spring	Landvik	Transition	15	15.0
57	WhiteSea	Onega Bay, White Sea	Pacific	30	5.0
58	KandalakshaBay_WhiteSea	Kandalaksha Bay, White Sea	Pacific	24	5.0
59	KandalakshaBay_WhiteSea	Kandalaksha Bay, White Sea	Pacific	19	10.0
60	PechoraSea_BarentsSea	Pechora Sea	Pacific	23	5.0

CNV=Average copy number

Supplementary Table 14: Chromatographic separation used for PRM LC-MS/MS analysis of sperm and tissue samples.

Time [min]	Flow [μ l/min]	% mobile phase B*
0	0.3	5
0.1	0.3	5
0.3	0.3	5
3.0	0.3	10
47.0	0.3	45
47.1	0.3	90
51.9	0.3	90
52.0	0.3	5
58.0	0.3	5

*See Methods for details