BMC Ecology and Evolution

RESEARCH

Open Access

Population stratification in Northern shrimp (*Pandalus borealis*) off Iceland evident from RADseq analysis

Áki Jarl Láruson[1*], Hafrún Gunnarsdóttir[2,3], Jake Goodall[3,4], Snæbjörn Pálsson[3] and Ingibjörg G. Jónsdóttir[3]

Abstract

The northern shrimp *Pandalus borealis* (ice. Stóri kampalampi) is a North Atlantic crustacean of significant commercial interest which has been harvested consistently in Icelandic waters since 1936. In Icelandic waters, the length at which this protandrous species transitions from male to female differs between the inshore and offshore populations, suggesting a biologically meaningful stratification which may or may not be plastic. Using reduced representative genomes assembled from RADseq data, sampled from 96 individuals collected at two time points (2018 and 2021), we compare the level of genetic structure across a gradient extending out of Skjálfandi bay, north Iceland. These data are compared to samples from a far offshore site, some 65 km out from the bay, as well as another inshore fjord in Arnarfjörður, in northwestern Iceland. Since 1999, no harvesting of inshore populations of *P. borealis* in Skjálfandi has been allowed due to stock decline, but harvesting of offshore stocks has continued. Uncertainty surrounding the extent of structure between the in- and offshore aggregations has remained. Here we report distinct genetic structure defining the inshore and offshore populations of northern shrimp, but find significant admixture between the two. Most importantly, we see that genetically inshore populations of northern shrimp extend far outside the harvest boundaries of inshore shrimp, and offshore individuals may exhibit punctuated migration into the inshore areas.

Keywords Northern shrimp, Iceland, Population genetics, *Pandalus borealis*, Molecular ecology, Seascape genetics

Background

The potential for connectivity is a uniquely challenging consideration in the study of marine population divergence. Dynamic ocean current systems, in conjunction with durable pelagic larval stages, can connect remote and seemingly disparate sites and permit population connectivity in spite of physiological and geographic expectations. In fact, speciation in spite of seemingly rare hard environmental barriers is a long standing theme in marine evolutionary studies [1, 2]. With a growing body of genome level analysis available for marine populations the mystery only seems to deepen. Low overall genomic divergence belies the separation in marine populations at distinct hot spots of accumulated differences [3–5].

*Correspondence:
Áki Jarl Láruson
aki.jarl.laruson@hafogvatn.is
[1]Marine and Freshwater Research Institute, Hafnarfjörður, Iceland
[2]Northeast Iceland Nature Research Centre, Húsavík, Iceland
[3]Faculty of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland
[4]Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

Understanding underlying genomic architecture is thus critical when considering population structure, as different parts of the genome can experience varying influences of connectivity and adaptation [6]. Within the marine environment, the gradient from inshore to offshore sites represents a consistent environmental feature that drives signatures of divergence [7–9]. Inshore and offshore environments demarcate significant ecological shifts in the marine environment, which affects both overall species composition, as well as diversity gradients within species that occupy both types of habitats [7, 10, 11].

The northern shrimp, *Pandalus borealis* (ice. Stóri kampalampi), is a commercially important North Atlantic caridean decapod. Across its broad boreal distribution, the northern shrimp is characterized by a lack of significant genetic structure [12–15]. A recent study of northern shrimp transcriptomes off the coast of southeastern Canada has suggested a plasticity (differential gene expression) in responding to habitat heterogeneity observed between populations [15]. However, allozyme analysis has suggested that inshore and offshore sites around Iceland may constitute possible population substructure [12]. Microsatellite analysis of Norwegian northern shrimp suggests a similar inshore and offshore dynamic between Norway's northernmost coast and the Barents sea, with no well defined barrier [8]. The species is protandrous (reaching sexual maturity first as a male before developing into a female past a certain body size), and assessments of the length at which individuals transition from male to female has been proposed to differ between inshore and offshore populations in northern Iceland, suggesting an important biological distinction between the two [16, 17]. Whether this phenotypic distinction is driven by distinct accumulations of genetic mutations or plastic differences remains to be definitively determined, but is partially addressed by the work presented here.

Harvesting of northern shrimp began in Iceland in 1936 in two fjords in the northwest of the island, Arnarfjörður (Fig. 1) and Ísafjarðardjúp. In 1990, Skjálfandi bay (Fig. 1) in the north was one of the last areas where shrimp were found in catchable quantities. Skjálfandi bay is a large glacially formed bay in northern Iceland and is characterized by a steep depth and salinity gradient as two large rivers, one glacial and another springfed, flow into the bay, providing a clear environmental cline. Shrimp fishing was closed after a large drop in the stock index of northern shrimp in Skjálfandi in 1999, and the index has remained low since then. Since the shrimp fishery closure, only two smaller scale harvests have taken place in 2012/13 and 2015/16 [18].

The inshore shrimp area at Skjálfandi is delimited by a line that runs across the bay, from the island of Flatey in the west to the Mánáreyar islands in the east (see blue line in Fig. 1). This line was recognized by the Icelandic parliament in 1997 to broadly define inshore and offshore fishing areas, with no specific consideration of northern

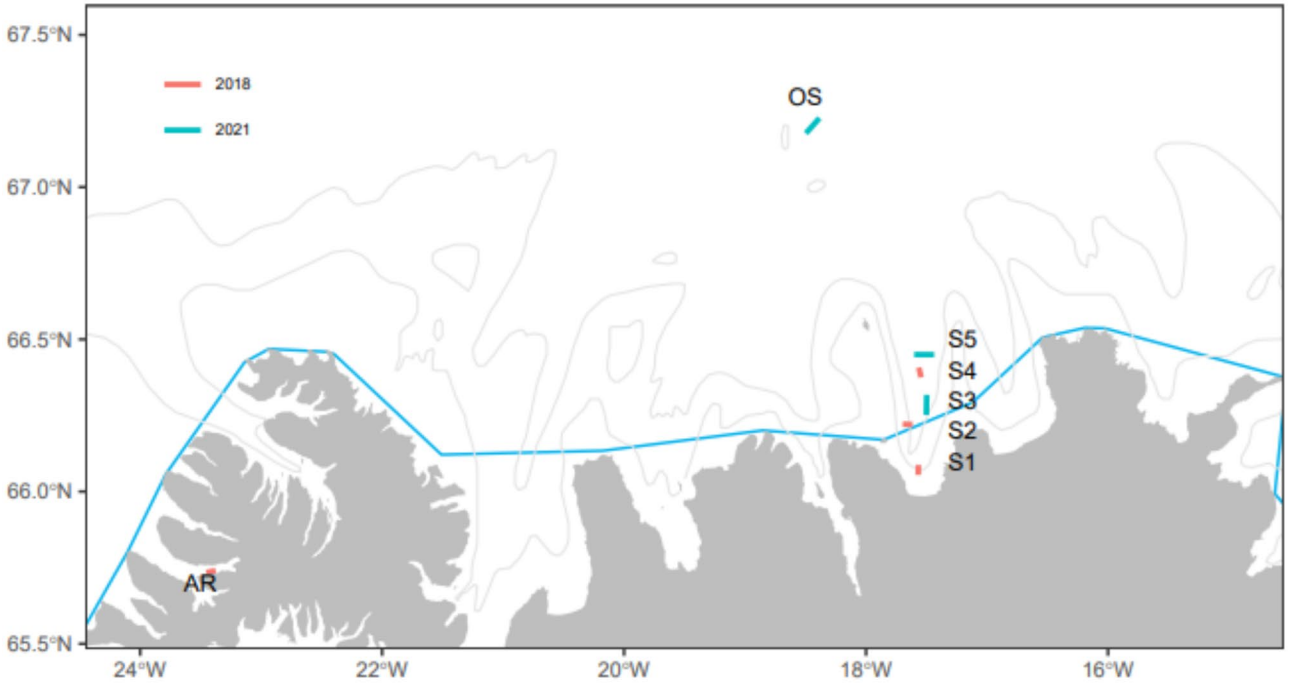

Fig. 1 Sites sampled during surveys in October 2018 (cyan) and March 2021 (orange). The blue line represents the current demarcation of inshore and offshore *P. borealis* stocks. AR = Arnarfjörður; S1-S5 = Skjálfandi (increasing numbers correspond to distance from shore); OS = Offshore site off Kolbeinsey

shrimp biology. Shrimp caught on the offshore side of the line belong to the offshore shrimp quota, while those caught inside the line belong to the inshore shrimp quota at Skjálfandi. Although shrimp are relatively local, post settlement, they show some movement and no hard geographical barriers are present to stop them from moving across the line that demarcates the zones [19]. The distinction between inshore and offshore northern shrimp is therefore not clear, and offshore shrimp have been caught as a part of the inshore shrimp harvest in Skjálfandi bay. There are no definitive studies on the extent to which these inshore and offshore clusters mix. However, life history and genetics studies using allozyme markers carried out between 1988 and 1995 indicated a potential contact zone in the North of Iceland, approximately 20–40 km offshore from the defined inshore shrimp harvest areas [17].

The aim of this study is to assess the genetic population structure of the northern shrimp in Skjálfandi bay, specifically across the inshore/offshore defined fisheries line, and characterize any structural patterns from the inshore to the offshore region north of Iceland in the context of a measurable phenotypic variation.

Methods

Sampling

A total of 96 samples were collected for genomic analysis in annual surveys in October 2018 and March 2021 (Fig. 1; Table 1). In October 2018, a total of 62 individuals were collected in the annual shrimp survey conducted by the Marine and Freshwater Research Institute (MFRI) in Arnarfjörður and Skjálfandi where the overall shrimp stock biomass is estimated to inform fisheries management. Some 46 individuals were collected from three stations along a transect from the innermost station in Skjálfandi towards an outer station located in the offshore shrimp area (Fig. 1: S1, S2, S4 - higher numbers indicating a greater distance from shore). An additional 16 individuals were collected in Arnarfjörður (AR), in north-west Iceland. In March 2021, a total of 34 individuals were collected in the annual groundfish bottom trawl survey conducted by MFRI. Samples were collected from three stations; 23 individuals from two stations along the transect where the 2018 samples were collected (Fig. 1: S3 & S5), and 11 individuals from an offshore site off Kolbeinsey (OS). No shrimp were caught within Skjálfandi bay in 2021.

All samples from both surveys consisted only of larger (female) shrimp. For each individual, a piece of muscle tissue ca. 0.2 gr from the tail was cut off and placed in a 2 ml tube containing 96% ethanol. Sea temperature was measured at each station using a pre-calibrated Scanmar clump sensor (SS4-C-VTLA 144) attached to the trawl headline (bottom temperature) or the vessel bottom (surface temperature). Salinity data for sample sites at similar times of year were collected from SeaDataNet (https://www.seadatanet.org).

Morphology

To assess phenotypic variability between our sampled sites, measurements from 250 individual shrimp caught per site per year during previous annual shrimp surveys (1988–2020) were compiled for the Skjálfandi sites S1 & S4, as well as the OS site. The proportion of females at carapace length p were fit with a generalized model (GLM) with a logit link function:

$$p = \frac{1}{1 + e^{-k(TL - L50)}}$$

where TL is the carapace length in mm, k is the slope of the logistic curve and L_{50} is the carapace length at which 50% of the biomass of caught shrimp were females (i.e. the size at sex change). The GLM model was used to estimate L_{50} for each site (S1, S3, & OS) across the survey years. Data from single stations in the annual offshore survey were used to estimate L_{50} for sites OS and S3, but data from four stations within the innermost part of Skjálfandi were used for the S1 site. A paired t-test comparing L50 data between individual sites was performed with the R stats package [20] function t.test.

Table 1 Number of individuals collected from each site, the year the samples were collected, mean latitude and longitude of collection trawl, and number of individuals retained for analysis post filtration. Mean values of environmental measurements (temperature, salinity, and depth) per site are also included

Site	Year	Lat.	Lon.	Ind. sampled	Ind. post filtration	Bottom Tmp. (°C)	Surface Tmp. (°C)	Salinity (*P.S.U.*)	Depth (m)
AR	2018	65.737	-23.4148	16	10	3.8	7.1	34.38	77.5
S1	2018	66.071	-17.5675	15	10	5.8	5.5	34.53	144.5
S2	2018	66.221	-17.6563	15	9	5.6	5.4	34.82	213
S3	2021	66.285	-17.4998	11	10	2.8	2.2	34.83	223.5
S4	2018	66.391	-17.5543	16	7	5.1	5.1	34.94	266.5
S5	2021	66.45	-17.5193	12	11	2.7	2.3	34.8	280
OS	2021	67.737	-18.4408	11	11	-0.1	2.0	34.74	458

DNA sequence generation

DNA was extracted using a Machery-Nagel NucleoSpin Tissue mini kit for DNA from cells and tissue (Item # 740952.50). Double digest RAD sequencing libraries were constructed as in Goodall et al. (2021) by digesting whole genomic DNA with *Sau*3AI and *Ape*KI and selecting fragments 390 bp to 430 bp using a Pippin Prep (Sage Science). Molecular identification tags were assigned to individuals using combinatorial barcoding of forward (5 bp) and reverse (6 bp) reads, respectively, and pooled into a library before size selection and amplification using 10 PCR cycles. Resulting library concentrations were quantified using SYBR Gold double-stranded DNA assay measured on a TECAN GENios plate reader (TECAN™, www.tecan.com), and the sizes estimated by 2% agarose gel. Libraries were sequenced at Novogene UK, using Illumina HiSeq 2500 (paired-end 2 × 125 bp) on a single lane, generating an average of 30.3 million paired raw reads per individual and with base error rate of 0.03%.

Variant site calling and filtering

Single nucleotide polymorphisms (SNPs) were called using Stacks v.2.0b [21, 22] with the following options enabled: -inline_inline,–r,–c,–q,–t = 119. Subsequent filtering was performed with VCFtools (0.1.16) [23]. Variants that were genotyped at fewer than 50% of sites across all individuals (--max-missing 0.5), had a minor allele count below three across all individuals (--mac 3), and had a sequencing depth less than 10 for each individual (–minDP 10), were removed. The amount of missing genotype calls per individual from the 2018 and the 2021 samples were quantified (Figure S1, Table S1), and subsequently individuals missing more than 50% of the polymorphic loci were removed. This reduced the number of individuals in the dataset from 96 to 68 (42% reduction in the 2018 batch, 6.25% reduction in the 2021 batch, Table 1). Following the reduction in the total number of individuals, sites with more than 5% of genotype calls missing (--max-missing 0.95) and a mean depth lower than 20 across all the remaining individuals (--minmeanDP 20) were then filtered out. Finally, sites that at this point had more than 10% missing genotype calls per sampling site were filtered. For population structure analysis, one random SNP per contig was retained via the bcftools prune plug in (-w 170 bp -n 1 -N rand) [24, 25].

Genetic analysis

Weir & Cockerham's (WC) F_{ST} values between all sampling sites were calculated with the weir-fst-pop command in VCFtools (0.1.16) [23] and bootstrap analysis of WC F_{ST} values was done with R package *StAMPP* [26]. Ordination analysis of allele frequencies was done with SMARTPCA, implemented in the EIGENSOFT (v.8.0.0) package [27, 28]. Distances were calculated from GPS sample coordinates with the arcdist function from the R package *geo* (v. 1.5-1) [29]. Isolation by distance was estimated via a Mantel test with a Spearman's correlation method, as implemented in the R package *vegan* (v.2.6-4) [30]. Additional summaries (H_O, H_S, H_T, F_{ST}, F_{IS}) were calculated with the R package *hierfstat* (v.0.5–11) [31].

The filtered VCF file was converted to a "12" coded. ped file via PLINK (v. 1.90b7.2) [32], before structure analysis was performed with the program ADMIXTURE (v1.3) [33]. A 100-fold cross validation was specified, for 200 iterations, for each proposed number of populations ranging from 1 to 15. Results were visualized in R using *ggplot2* (v.3.5.1) [34].

Outlier detection was done using two different approaches. First via the PCA based R package *PCAdapt* (v. 4.4.0) [35], starting with an initial spot check of 20 principal components, before limiting the analysis to the first two principal components based on the percentage of variance explained. Converted q-values were bounded by a 5% false discovery rate. A minimal minor allele frequency threshold of 0.1% was used, as the input data was already heavily filtered. Finally, Bayescan (v. 2.1) [36–38] was used with the -n 5000, -burn 50,000, and -all_trace -pr_odds 1000 parameter specification. VCF to geste file format conversion was done with PGDSpider (v. 3.0.0.0) [39]. A redundancy analysis of outlier loci association to measured environmental variables was done with the *vegan* R package (v.2.6-4) [30].

Results

A clear distinction in the size at which northern shrimp transition from male to female was seen between the Skjálfandi sites and the offshore (OS) site (Fig. 2). All individuals sampled for DNA reflected the typical size distribution of the female shrimp previously caught across the sample sites. Despite yearly fluctuations in the length at which 50% of the caught shrimp had transitioned to the terminal female stage, female shrimp caught at the OS site were on average 28% longer than at either of the Skjálfandi sites. A paired t-test showed minor difference between the S1 and S2 sites L50 data ($t = 3.99$, p-value = 0.0006), but showed OS site data to be significantly longer than either the S1 ($t = 22.39$, p-value $< 2.2 \times 10^{-16}$) or S4 data ($t = 21.96$, p-value $< 2.2 \times 10^{-16}$).

A total of 419,723 polymorphic loci were identified by Stacks across 96 individuals. There was a clear batch effect of the timing of the sampling and the number of missing loci per individual (Table S1). While both the 2018 and 2021 samples were kept frozen until they were extracted and sequenced in the same batch, the older samples had on average over 85% more missing data (Figure S1 A). Filtration at the allele, individual, and sample site level left 2,364 sites across 68 individuals that were

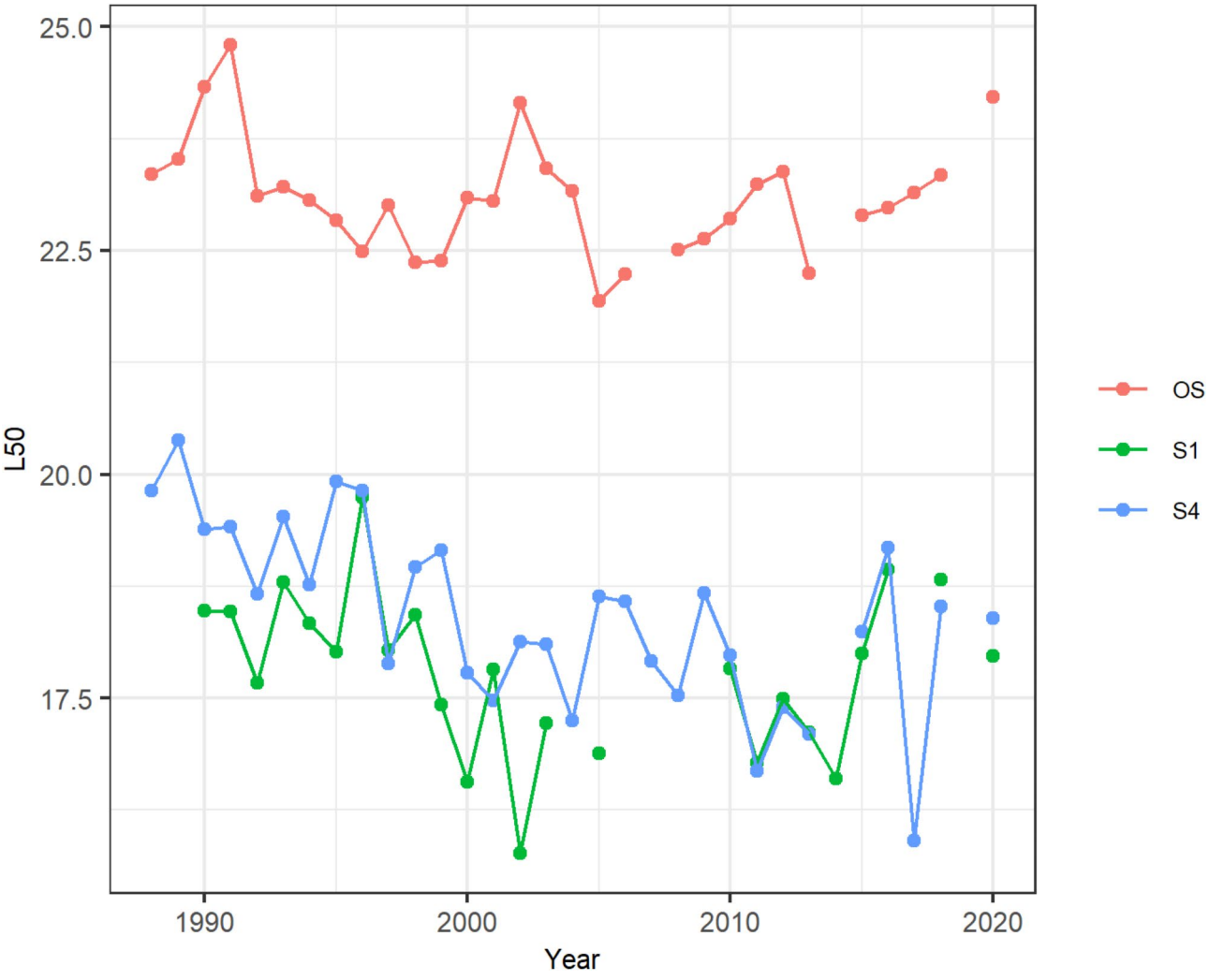

Fig. 2 Measures of length at which 50% of surveyed northern shrimp biomass was female. At each survey site for each year the biomass is based on 250 individual northern shrimps. The OS site consistently shows that terminal phase females are longer than at either the innermost site in Skjálfandi bay (S1) or a site further out of the bay, well past the inshore/offshore delimiting harvest line (S4)

Table 2 Weir & Cockerham's F_{ST} values between all sampling sites listed below the diagonal. P-values from 1000 bootstrap resamples are listed above the diagonal. Highest pairwise F_{ST} values are in bold

F_{ST}	S1	S2	S3	S4	S5	OS	AR
S1		0.052	0.066	0.014	0.391	0.000	0.058
S2	0.006	NA	0.005	0.203	0.008	0.000	0.282
S3	0.005	0.009	NA	0.078	0.153	0.000	0.024
S4	0.008	0.004	0.006	NA	0.018	0.000	0.384
S5	0.001	0.007	0.003	0.007	NA	0.000	0.032
OS	**0.025**	**0.045**	**0.029**	**0.042**	**0.013**	NA	0.000
AR	0.006	0.002	0.007	0.001	0.006	**0.033**	NA

representative across both the 2018 and the 2021 samples (Figure S1 B) that were used for outlier detection. Following filtration for one random SNP per contig 1,099 SNPs remained for population structure analysis, with a mean sequencing depth of 106.5X across all individuals (94X median depth).

Genome level divergence, as assessed via F_{ST} between the different sample sites, was generally low (F_{ST}<0.0065) with the notable exception of the offshore site compared to all other sites within Skjálfandi bay and Arnarfjörður (mean F_{ST} = 0.027, Table 2). Only the F_{ST} values which included comparisons with the OS site were significant

after adjusting the p-value significance threshold due to multiple tests.

Scaled F_{ST} distances from offshore towards land in Skjálfandi bay did not significantly increase with geographic distance (Fig. 3). A Mantel test showed no isolation by distance (significant relationship between F_{ST} and distance between the inner bay and the offshore site) ($r = 0.589$, p-value = 0.133).

Ordination analysis from smartPCA showed the biggest overall drop in Eigenvalue following PC1 (Figure S2). In line with the result from the pairwise F_{ST} comparisons, the PCA results showed a high level of overlap between inshore sample sites across Skjálfandi and the Arnarfjörður samples (Fig. 4). The offshore site clustered separately from all other samples, with a slight overlap with the site furthest outside of Skjálfandi bay (site S5) and a single individual from the site furthest into the bay (site S1).

Similarly, a cross validation error analysis from the ADMIXTURE results indicated that a K = 2 best fit the data. This split the OS site from the remainder of the samples as a distinct ancestry assignment. All other sites contained signatures of offshore admixture and the innermost site within Skjálfandi (S1) contained an individual fully assigned to an offshore ancestry. Inshore individuals with no signature of admixture, i.e., with fully inshore ancestry, were in the minority at every site except S2 (5 of 9). Looking at the ancestry assignment to the OS group (blue color in Fig. 5), we see a general pattern of increasing average individual assignment as you move out of the Skjálfandi bay towards the offshore site. The same pattern of a K = 2 admixture was observed when considering the 2018 and the 2021 samples separately.

Some 571 alleles were identified as being potentially under balancing selection by Bayescan, with no loci suggested to be under directional selection. A total of 15 alleles were identified as potentially being under selection by *PCAdapt*, and nine of those alleles were also identified by Bayescan. Those nine alleles appear to fluctuate in frequencies in the inshore environments but tend towards elevated or lower frequencies at the OS site (Figure S3). None of the nine alleles had elevated sequencing depth, as might be expected from paralogous sequences that might erroneously signal balancing selection (mean 56.2X, min 25X, max 94X). Five of those nine alleles gave negative F_{IS} values (Table 3), suggesting a greater observed heterozygosity than expected. Averaged within site F_{IS} values were all positive (Table S2). Redundancy

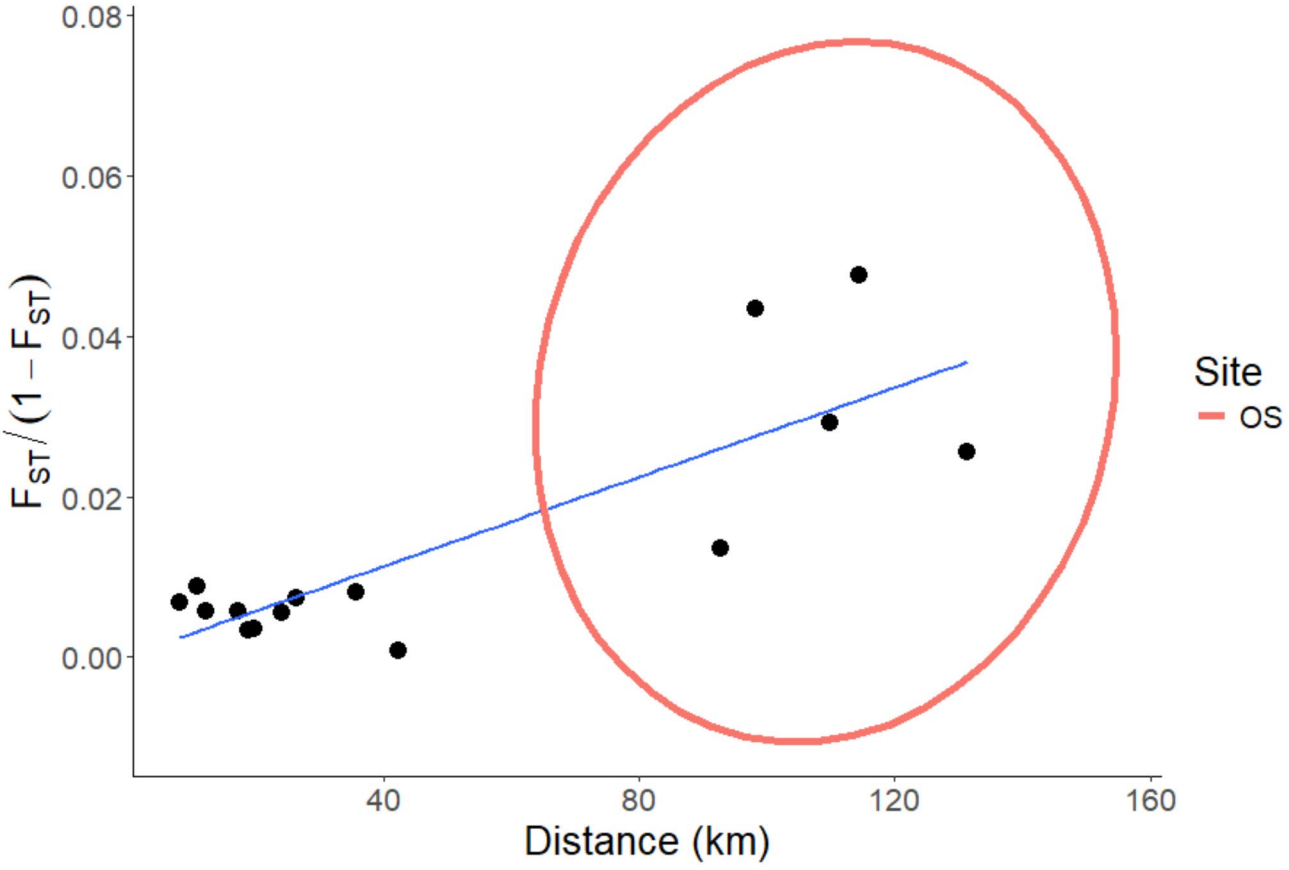

Fig. 3 Geographical distance between sites plotted against a corrected F_{ST} calculation for those same two sites. The points circled in red are all the F_{ST} values between the offshore site (OS) and the Skjálfandi sites (S1, S2, S3, S4, & S5)

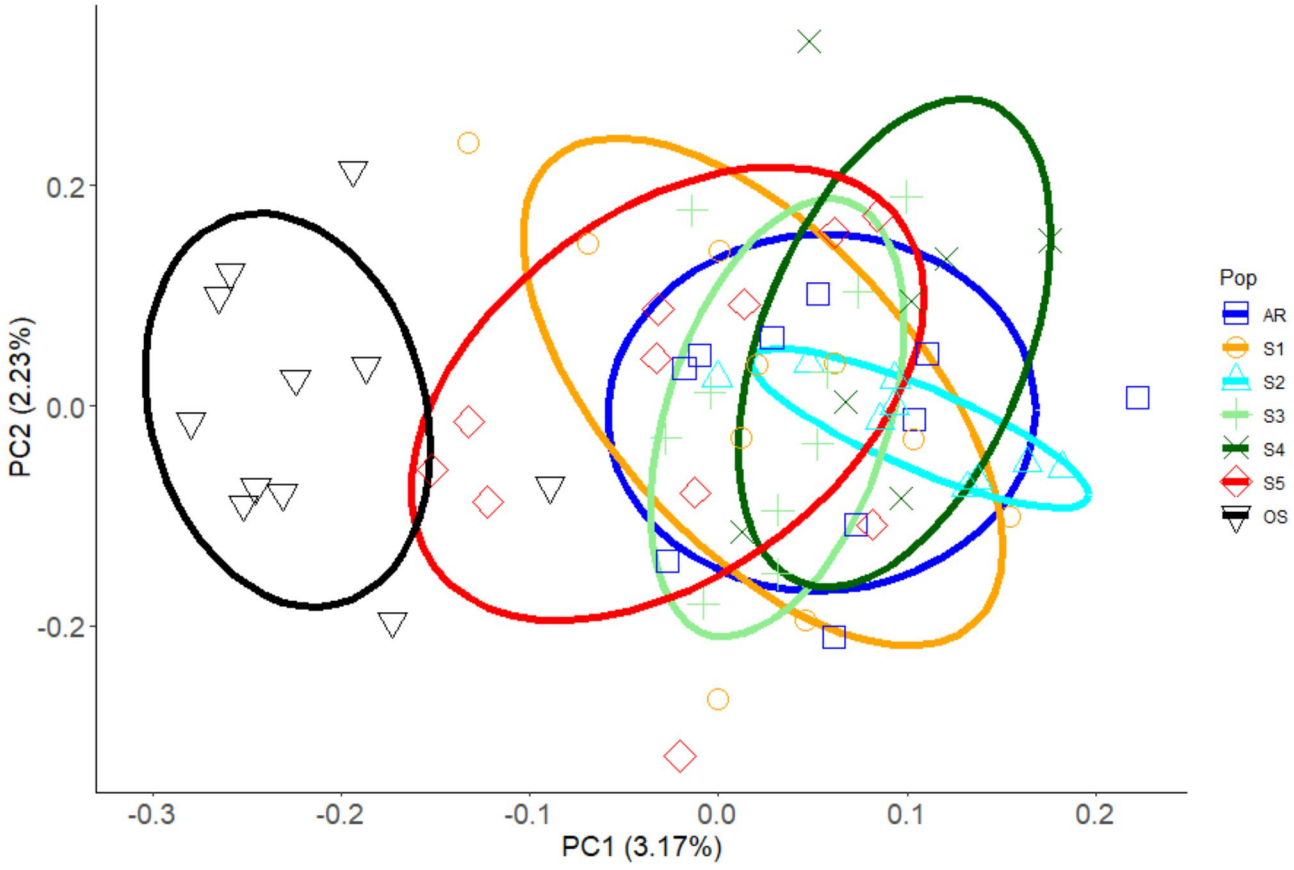

Fig. 4 Results of PCA ordination of allele frequencies, performed with the smartPCA algorithm. The distinction between the offshore site (OS, outlined in black) and the remaining samples in the study appears to drive the variation observed across the first PC. While there is significant overlap between all the remaining sites, three individuals from the outermost Skjálfandi bay site (S5) and one from the innermost site (S1) appear to cluster closer to the OS samples

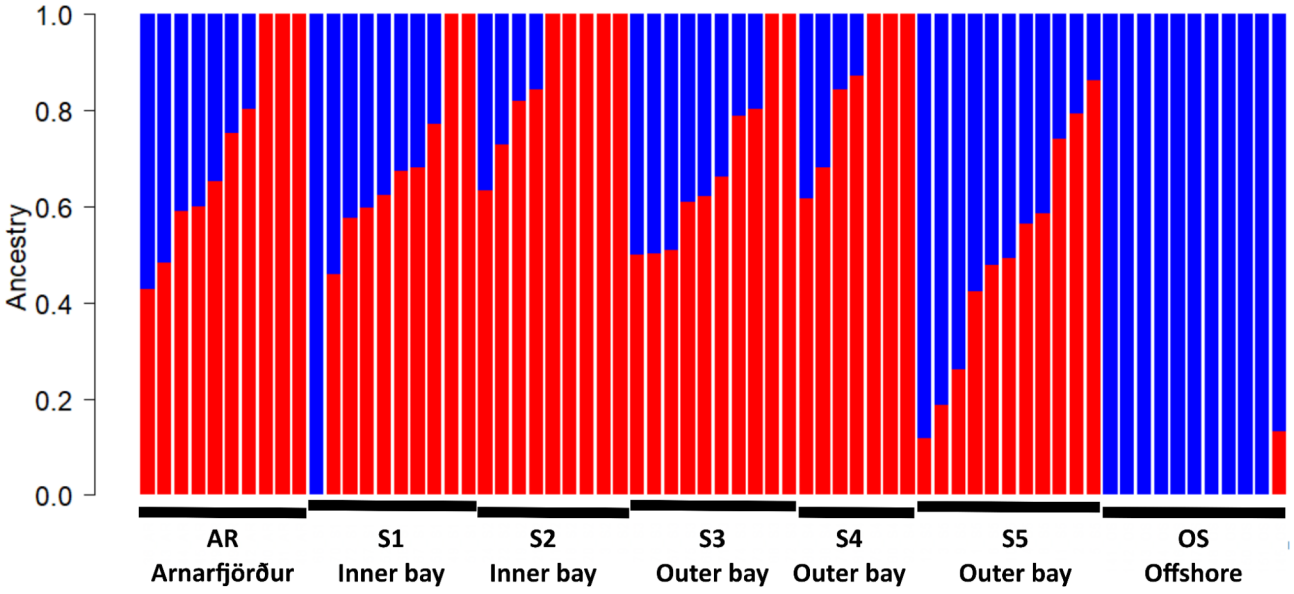

Fig. 5 Ancestry assignment (Q values) for k = 2 from ADMIXTURE, plotted across sampled sites. Cross validation error values for k values immediately above 2 were larger and are not shown here. The two populations identified conform to an inshore (red) and an offshore (blue) assignment. One individual at the innermost site of Skjálfandi bay (S1) appears to be an unambiguous offshore northern shrimp, while three individuals at the outermost site (S5) have a majority offshore assignment

Table 3 Observed heterozygosity (H_O), within population genetic diversity (H_S, expected heterozygosity), total genetic diversity (H_T), F_{ST}, and F_{IS} for the nine outlier alleles

Allele	HO	HS	HT	FST	FIS
X34675_43	0.4356	0.3664	0.3814	0.0393	-0.1890
X10400_73	0.4223	0.3623	0.4118	0.1203	-0.1656
X3587_34	0.4078	0.3601	0.4178	0.1381	-0.1325
X9542_61	0.3902	0.3466	0.3992	0.1319	-0.1259
X27061_111	0.4327	0.3979	0.3990	0.0029	-0.0877
X24089_104	0.4265	0.4316	0.4950	0.1279	0.0119
X8114_74	0.3383	0.3672	0.4289	0.1439	0.0787
X20170_116	0.3598	0.4567	0.4969	0.0808	0.2123
X28654_117	0.0130	0.0447	0.0443	-0.0096	0.7098

analysis including bottom temperature, surface temperature, and salinity (depth was not included due to high correlation to the other three variables) showed that over 55% of the variation in outlier allele frequency could be explained by bottom temperature alone, compared with less than 68% explained when surface temperature and salinity were included (Figure S4).

Discussion

From two sampling efforts in the fall of 2018 and spring of 2021, we generated a reduced representative genomic dataset from two purported *P. borealis* inshore shrimp fisheries in the West (Arnarfjörður) and North (Skjálfandi) of Iceland, as well as from an offshore site some 130 km North of Skjálfandi bay. Interestingly, northern shrimp from the two sites Skjálfandi and Arnarfjörður, which are some 290 km apart, generally showed genomic divergence levels on a par or lower to those seen between sites within the northern bay Skjálfandi (less than a 45 km maximum site distance). The level of divergence seen from F_{ST}, PCA, and admixture analysis of the offshore site from all inshore sites identifies a genetically distinct population split between inshore and offshore northern shrimp around Iceland. These results do suggest that the present line demarcating the inshore and offshore harvest quotas, which designates only the sites AR and S1 as inshore while all others are defined as offshore (Fig. 1), does not reflect a realistic population demarcation.

This study reaffirms patterns described by Jónsdóttir et al. (1998) where a genetic distinction between offshore and inshore shrimp samples is observed. However this signature is confounded by a signature of potentially unidirectional gene flow from the offshore to the inshore. A minority of individuals were classifiable as clearly inshore, and only samples significantly offshore (OS site, over 130 km out from the inner bay) showed little to no signature of admixture. While seasonal differences between the 2018 (October) and 2021 (March) sampling events potentially influenced the individual levels of admixture observed at any given inner or outer bay site, it seems that admixture between these populations is a consistent occurrence. Analysis of the 2018 samples, which were taken in October, still supports a signal of binary admixture across all sites (all 2018 sites are inshore sites), and similarly 2021 sites (sampled in March) analyzed separately show admixture in the two inshore sites and an homogenous "offshore" assignment of the OS site.

A recent study of northern shrimp in southeast Canada [15] showed a very modest, but significant, genetic divergence (as measured by WC F_{ST}) of an offshore sample from an Eastern Scotian Shelf site from two more inshore sites (St. Lawrence Estuary & Northeast Newfoundland Coast) despite being geographically closer to either inshore site than the two inshore sites were to each other. However this same study revealed a significant capacity for transcription regulation in northern shrimp in response to changing temperatures, which did not appear to be variable based on local conditions (i.e., locally adapted). The authors concluded that given the overall low levels of genetic diversity observed, despite the broad and variable habitat distribution of northern shrimp, the species is dependent on transcriptional plasticity to endure habitat variation. This limited adaptive potential was suggested to leave the species vulnerable to rapid environmental shifts beyond the plasticity threshold, especially at a large geographic scale.

However, given the signature of admixture between genetically distinguishable inshore and offshore populations in northern Iceland, it is possible that standing genetic variation in the species is enriched in select localities. Given the results of the redundancy analysis, we can assume that bottom temperature is a factor influencing the allele frequencies of identified outlier alleles potentially under balancing selection. Temperature is a key factor in development and reproduction for northern shrimp [19], and therefore the outsized effect of bottom temperature on the variance in outlier allele frequencies suggests a level of temperature driven selection across the inshore and offshore populations. Given that, it is possible that the measurable phenotypic difference observed between the inshore and offshore locations (size at which individual shrimp transition from male to female) is not a fully plastic response as potentially presumed given the findings in Leung et al. (2023), but at least in part a genetically determined adaptation.

An important oceanographic feature of Skjálfandi bay is that during the winter months, arctic waters are carried south into the bay, potentially carrying offshore northern shrimp inshore and promoting admixture. This punctuated dynamic may be driving the signature of balancing selection on select alleles that allow for the "warmer" bottom temperature adapted inshore population to receive punctuated "colder" bottom temperature adapted alleles.

If this localized dynamic is promoting a greater standing genetic variation in local populations of northern shrimp, it may prove to be important for the adaptability of northern shrimp around Iceland as the ocean climate at northern latitudes continues to shift [40].

It should however be noted that signatures of balancing selection can be confounded by paralogous genes, which can inflate estimates of heterozygosity if not distinguishable in the sequence data [41]. Although we do not observe exceptionally high coverage of the outlier alleles identified in this study, higher coverage genome analysis across a broader geographic range (i.e. across multiple bays) should give a better insight into the role of environmentally associated alleles in demarcating the inshore and offshore populations of northern shrimp around Iceland.

Notably, the data presented here is RADseq reduced representative genomic data, meaning that genomic "hot spots" of divergence or local adaptation are likely to have been missed [42]. Even so, there does appear to be a suggestion of a temperature driven structure in this system. Similar patterns of a seemingly unidirectional offshore-to-inshore dynamics have also been seen in a diverse set of taxa around Iceland, such as in the Common Whelk (*Buccinum undatum*) [7], potentially in Cod (*Gadus morhua*) [9], and Sugar Kelp (*Saccharina latissima*, Láruson et al. *in review*).

Conclusion

We observe a genome wide distinction between inshore and offshore populations of northern shrimp in northern Iceland. The presence of offshore individuals within the inshore northern shrimp population suggest that there is not a clearly demarcated hybridization zone between the inshore to offshore sites, but a persistent primarily unidirectional gene flow from offshore to inshore. Given this dynamic, the present line demarcating the inshore and offshore harvest quotas of *P. borealis* is not reflecting the biological reality of the population structure. Due to the observed gradient from inshore towards offshore shrimp, with shrimp seasonally moving between the two legally defined areas, it is not possible to produce a fixed line that is biologically relevant.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12862-025-02369-9.

Supplementary Material 1

Acknowledgements
Salinity data were provided through SeaDataNet Pan-European infrastructure for ocean and marine data management (https://www.seadatanet.org).

Author contributions
IGJ conceived the research; IGJ & SP designed data acquisition methodology; HG, JG, & ÁJL contributed to data acquisition; SP & ÁJL analyzed the DNA data; IGJ & ÁJL wrote the first draft of the paper; all authors read, commented, and approved of the final manuscript.

Funding
All funding for this research was provided by the Marine & Freshwater Research Institute of Iceland.

Data availability
All unfiltered DNA sequence data has been deposited to the NCBI Sequence Read Archive (SRA) and are publicly available (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1200262). All R code used in the analysis is available on https://github.com/Hafro/P.borealis_RADseq.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 12 December 2024 / Accepted: 1 April 2025
Published online: 11 April 2025

References

1. Palumbi SR. Marine speciation on a small planet. Trends Ecol Evol. 1992;7(4):114–8. https://doi.org/10.1016/0169-5347(92)90144-Z.
2. Faria R, Johannesson K, Stankowski S. Speciation in marine environments: diving under the surface. J Evol Biol. 2021;34(1):4–15. https://doi.org/10.1111/jeb.13756.
3. Han F, Jamsandekar M, Pettersson ME, Su L, Fuentes-Pardo AP, Davis BW, Bekkevold D, Berg F, Casini M, Dahle G, Farrell ED, Folkvord A, Andersson L. Ecological adaptation in Atlantic herring is associated with large shifts in allele frequencies at hundreds of loci. ELife. 2020;9:e61076. https://doi.org/10.7554/eLife.61076.
4. Sanford E, Kelly MW. Local adaptation in marine invertebrates. Annual Rev Mar Sci. 2011;3:509–35. https://doi.org/10.1146/annurev-marine-120709-142756.
5. Tigano A, Jacobs A, Wilder AP, Nand A, Zhan Y, Dekker J, Therkildsen NO. Chromosome-level assembly of the Atlantic silverside genome reveals extreme levels of sequence diversity and structural genetic variation. Genome Biol Evol. 2021;13(6):evab098. https://doi.org/10.1093/gbe/evab098.
6. Akopyan M, Tigano A, Jacobs A, Wilder AP, Baumann H, Therkildsen NO. Comparative linkage mapping uncovers recombination suppression across massive chromosomal inversions associated with local adaptation in Atlantic silversides. Mol Ecol. 2022;31(12):3323–41. https://doi.org/10.1111/mec.16472.
7. Goodall J, Westfall KM, Magnúsdóttir H, Pálsson S, Örnólfsdóttir EB, Jónsson ZO. RAD sequencing of common Whelk, *Buccinum undatum*, reveals fine-scale population structuring in Europe and cryptic speciation within the North Atlantic. Ecol Evol. 2021;11:2616–29. https://doi.org/10.1002/ece3.7219.
8. Hansen A, Westgaard J-I, Søvik G, Hanebrekke T, Nilssen EM, Jorde PE, Albretsen J, Johansen T. Genetic differentiation between inshore and offshore populations of Northern shrimp (Pandalus borealis). ICES J Mar Sci. 2021;78(9):3135–46. https://doi.org/10.1093/icesjms/fsab181.
9. Pampoulie C, Berg PR, Jentoft S. Hidden but revealed: after years of genetic studies behavioural monitoring combined with genomics uncover new insight into the population dynamics of Atlantic Cod in Icelandic waters. Evol Appl. 2023;16:223–33. https://doi.org/10.1111/eva.13471.
10. Hoelzel AR, Potter CW, Best PB. Genetic differentiation between parapatric 'nearshore' and 'offshore' populations of the bottlenose Dolphin. Proc Royal Soc B. 1998;265(1177–1183). https://doi.org/10.1098/rspb.1998.0416.

11. Taillebois L, Barton DP, Crook DA, Saunders T, Taylor J, Hearnden M, Saunders RJ, Newman SJ, Travers MJ, Welch DJ, Greig A, Dudgeon C, Maher S, Ovenden JR. Strong population structure deduced from genetics, otolith chemistry and parasite abundances explains vulnerability to localized fishery collapse in a large Sciaenid fish, *Protonibea diacanthus*. Evol Appl. 2017;10:978–93. https://doi.org/10.1111/eva.12499.
12. Jónsdóttir ÓDB, Imsland AK, Nævdal G. Population genetic studies of Northern shrimp, *Pandalus borealis*, in Icelandic waters and the Denmark Strait. Can J Fish Aquat Sci. 1998;55(3):770–80.
13. Martinez I, Aschan M, Skjerdal T, Aljanabi SM. The genetic structure of *Pandalus borealis* in the Northeast Atlantic determined by RAPD analysis. ICES J Mar Sci. 2006;63(5):840–50. https://doi.org/10.1016/j.icesjms.2006.03.006.
14. Jorde PE, Søvik G, Westgaard J-I, Albretsen J, André C, Hvingel C, Johansen T, Sandvik AD, Kingsley M, Jørstad KE. Genetically distinct populations of Northern shrimp, *Pandalus borealis*, in the North Atlantic: adaptation to different temperatures as an isolation factor. Mol Ecol. 2015;24:1742–57. https://doi.org/10.1111/mec.13158.
15. Leung C, Guscelli E, Chabot D, Bourret A, Calosi P, Parent GJ. The lack of genetic variation underlying thermal transcriptomic plasticity suggests limited adaptability of the Northern shrimp, *Pandalus borealis*. Front Ecol Evol. 2023;11. https://doi.org/10.3389/fevo.2023.1125134.
16. Skúladóttir U. Size at sexual maturity of female Northern shrimp (*Pandalus borealis* Krøyer) in the Denmark Strait 1985–93 and a comparison with the nearest Icelandic shrimp populations. J Northwest Atl Fisheries Sci. 1998;24:27–37.
17. Skúladóttir U, Pétursson G. Defining populations of Northern shrimp, *Pandalus borealis* (Krøyer, 1838), in Icelandic waters using the maximum length and maturity of females. Rit Fiskideildar. 1999;16:247–62.
18. Marine and Freshwater Research Institute. (2020) State of Marine Stocks and Advice: Northern shrimp (Skjálfandi). Hanfarfjörður, Iceland. https://www.hafogvatn.is/static/extras/images/041-raekja_skjalf1225305.pdf
19. Skúladóttir U, Jónsson E. Rækjan Við Ísland. Ægir. 1980;2:86–95117.
20. R Core Team. (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
21. Catchen J, Amores A, Hohenlohe P, Cresko W, Postlethwait J. Stacks: Building and genotyping loci de Novo from short-read sequences. G3: Genes Genomes Genet. 2011;1:171–82.
22. Catchen J, Hohenlohe P, Bassham S, Amores A, Cresko W. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22:3124–40.
23. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. 1000 Genomes project analysis group, the variant call format and vcftools. Bioinformatics. 2011;27(15):2156–8. https://doi.org/10.1093/bioinformatics/btr330.
24. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map (SAM) format and SAMtools. Bioinformatics. 2009;25:2078–9. 1000 Genome Project Data Processing Subgroup.
25. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollyard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. Twelve years of samtools and BCFtools. GigaScience. 2021;10:giab008.
26. Pembleton L, Cogan N, Forster J. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour. 2013;13:946–52. https://doi.org/10.1111/1755-0998.12129.
27. Price A, Patterson N, Plenge R, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38:904–9. https://doi.org/10.1038/ng1847.
28. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006;2:12: e190. https://doi.org/10.1371/journal.pgen.0020190.
29. Björnsson H, Jónsson ST, Magnusson A, Elvarsson BT. (2022) geo: Draw and Annotate Maps, Especially Charts of the North Atlantic. R package version 1.5-1. https://github.com/hafro/geo
30. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Cunha R, Smith E, Stier T, Braak AT, C., and, Weedon J. (2022) vegan: Community Ecology Package. R package version 2.6-4 https://CRAN.R-project.org/package=vegan.
31. Goudet J, Jombart T. (2022) Hierfstat: Estimation and tests of hierarchical F-Statistics. R package version 0.5–11. https://CRAN.R-project.org/package=hierfstat
32. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, de Bakker PIW, Daly MJ, Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analysis. Am J Hum Genet. 2007;81(3):559–75.
33. Alexander DH, Novembre J, Lange K. Fast model-based Estimation of ancestry in unrelated individuals. Genome Res. 2009;19:1655–64.
34. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-; 2016.
35. Privé F, Luu K, Vilhjálmsson BJ, Blum MGB. Performing highly efficient genome scans for local adaptation with R package Pcadapt version 4. Mol Biol Evol. 2020;37(7):2153–4. https://doi.org/10.1093/molbev/msaa053.
36. Foll M, Gaggiotti OE. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. Genetics. 2008;180:977–93. https://doi.org/10.1534/genetics.108.092221.
37. Foll M, Fischer MC, Heckel G, Excoffier L. Estimating population structure from AFLP amplification intensity. Mol Ecol. 2010;19:4638–47. https://doi.org/10.1111/j.1365-294x.2010.04820.x.
38. Fischer MC, Foll M, Excoffier L, Heckel G. Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (Microtus arvalis). Mol Ecol. 2011;20:1450–62. https://doi.org/10.1111/j.1365-294x.2011.05015.x.
39. Lischer HEL, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 2012;28:298–9. https://doi.org/10.1093/bioinformatics/btr642.
40. Lotterhos KE, Láruson ÁJ, Jiang LQ. Novel and disappearing climates in the global surface ocean from 1800 to 2100. Sci Rep. 2021;11:15535. https://doi.org/10.1038/s41598-021-94872-4.
41. Hasselmann M, Vekemans X, Pflugfelder J, Koeniger N, Koeniger G, Tingek S, Beye M. Evidence for convergent nucleotide evolution and high allelic turnover rates at the complementary sex determiner gene of Western and Asian honeybees. Mol Biol Evol. 2008;25(4):696–708. https://doi.org/10.1093/molbev/msn011.
42. Lowry DB, Hoban S, Kelley JL, Lotterhos KE, Reed LK, Antolin MF, Storfer A. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Mol Ecol Resour. 2017;17:142–52. https://doi.org/10.1111/1755-0998.12635.

Publisher's note