ICES Journal of Marine Science, 2023, **80**, 122–132
DOI: 10.1093/icesjms/fsac217
Advance access publication date: 5 December 2022
Original Article

ICES CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Two distinct population clusters of northern sand lance (*Ammodytes dubius*) on the northwest Atlantic shelf revealed by whole genome sequencing

Lucas F. Jones[1], R. Nicolas Lou[2], Christopher S. Murray[3], Dominique Robert[4], Christina M. Bourne[5], Caroline Bouchard[6], Marek Kučka[7], Y. Frank Chan[7], David B. Carlon[8], David N. Wiley[9], Nina O. Therkildsen[2] and Hannes Baumann[1,*]

[1]Department of Marine Sciences, University of Connecticut, 1080 Shennecossett Road, Groton, CT 06340, USA
[2]Department of Natural Resources and the Environment, Cornell University, 208 Fernow Hall, NY 14853, USA
[3]Biology Department, MS-32, Woods Hole Oceanographic Institution, MA 02543, USA
[4]Institut des sciences de la mer, Université du Québec à Rimouski, CP 3300, Rimouski, QC G5L 3A1, Canada
[5]Fisheries and Oceans Canada, Northwest Atlantic Fisheries Centre, St.John's, NL A1C 5X1, Canada
[6]Greenland Climate Research Centre, Greenland Institute of Natural Resources, Kivioq 2, P.O. Box 570, 3900 Nuuk, Greenland
[7]Friedrich Miescher Laboratory, Max Planck Society, 72076 Tübingen, Germany
[8]Department of Biology and Schiller Coastal Studies Center, Bowdoin College, Brunswick, ME 04011, USA
[9]NOAA/Stellwagen Bank National Marine Sanctuary, 175 Edward Foster Road, Scituate, MA 02066, USA

[*]*Corresponding author: tel:*+860-405-9297*;* e-mail: hannes.baumann@uconn.edu.

Northern sand lance (*Ammodytes dubius*) are essential forage fish in most offshore, temperate-to-polar waters on the Northwest Atlantic shelf (NWA), but their population structure and genetic separation from the American sand lance (*A. americanus*) remain unresolved. We assembled a reference genome for *A. dubius* (first in the Ammodytidae) and then used low-coverage whole genome sequencing on 262 specimens collected across the species distribution (Mid-Atlantic Bight to Greenland) to quantify genetic differentiation between geographic regions based on single nucleotide polymorphisms. We found strong separation between *A. dubius* from locations north and south of the Scotian Shelf, largely due to massive genetic differentiation spanning most of chromosomes 21 and 24. Genetic distance increased with geographic distance in the smaller southern cluster but not in the larger northern cluster, where genetic homogeneity appeared across large geographic distances ($>10^3$ km). The two genetic clusters coincide with a clear break in winter sea surface temperature, suggesting that differential offspring survival, rather than limited transport, causes a break in realized connectivity. Nuclear and mitochondrial DNA both clearly delineated *A. dubius* from *A. americanus*, thereby confirming a species boundary through spatial niche partitioning into inshore (*A. americanus*) and offshore (*A. dubius*) sand lance species on the NWA.

Keywords: *Ammodytes americanus*, *Ammodytes dubius* reference genome, *Ammodytes personatus*, connectivity, low-coverage whole genome sequencing, Scotian Shelf genetic break, Stellwagen Bank.

Introduction

Sand lances, or "sandeels" (Ammodytidae), are critically important forage fishes in temperate to polar shelf ecosystems across the northern hemisphere (Staudinger *et al.*, 2020). Their exceptional nutritional quality and high local abundance sustain commercial fish stocks (e.g. cod, tuna) and often promote spatial aggregations of iconic piscivores such as sharks, seabirds, and marine mammals (Engelhard *et al.*, 2014; Silva *et al.*, 2020). Some sand lance species are targeted commercially for oil and meal (Lynam *et al.*, 2013), as for example in the North Sea, where peak landings of >1 million tons annually (van Deurs *et al.*, 2009; ICES, 2010) caused declines of seabird and marine mammal populations in the late 1990s (Furness, 1999; Furness and Tasker, 2000; Daunt *et al.*, 2008). These cascading impacts indicate that sand lances comprise key components of any ecosystem-based management, which in turn requires better data and knowledge than we currently have for most species in this group. Sand lances are generally not represented well by fishery-independent surveys because their small, eel-like bodies pass through most trawl meshes, and—unlike any other fish group—they alternate between a pelagic feeding and a buried occurrence on diel and seasonal timescales (van Deurs *et al.*, 2009; Morrison, 2021). The burying behaviour constrains sand lances to specific, coarse-grain sand banks that produce characteristically patchy and temporally dynamic distributions (Richardson *et al.*, 2014; Wright *et al.*, 2019), but the connectivity among patches remains poorly understood.

On the Northwest Atlantic Shelf (NWA), the northern sand lance (*Ammodytes dubius*, Reinhardt, 1837) inhabits offshore sand banks from the Mid-Atlantic Bight to Labrador and Greenland (Figure 1; Robards *et al.*, 2000). Underscoring its ecosystem importance, known predators of *A. dubius* include 72 species of fish, squid, seabirds, and marine mammals (Staudinger *et al.*, 2020). The species has not been targeted commercially, and many basic aspects of its ecology, life history, climate sensitivity, and stock structure remain insufficiently studied. These knowledge gaps have inspired renewed

Received: August 1, 2022. **Revised:** November 6, 2022. **Accepted:** November 10, 2022

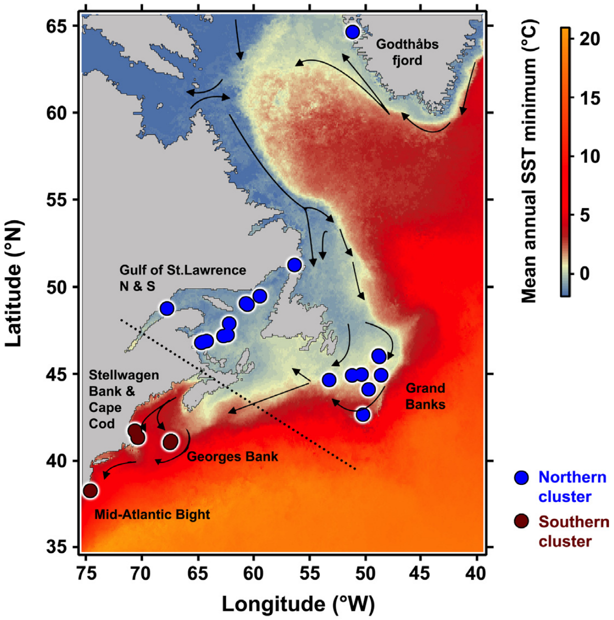

Figure 1. Map of the Northwest Atlantic (NWA) with regions and sampling sites for northern sand lance (*A. dubius*; see also Table 1). Red circles denote locations comprising the southern population cluster, while blue circles denote locations comprising the northern populations cluster, as revealed in this study. Contours depict the mean annual SST minimum (2000–2014) as a proxy for winter temperature (data from Bio-Oracle 2.0; Assis *et al.*, 2018), while black arrows illustrate surface currents. The dotted line visualizes the genetic break between northern and southern *A. dubius* population clusters somewhere on the Scotian Shelf.

research efforts in recent years, which have focused on sand lance from the southern Gulf of Maine (e.g. Stellwagen Bank), Cape Cod, and Nantucket Shoals (Silva *et al.*, 2020; Suca *et al.*, 2021; Baumann *et al.*, 2022; Suca *et al.*, 2022). What remains unresolved, but has clear conservation implications, is the genetic structure and population connectivity of *A. dubius* across its entire geographic distribution on the NWA. One hypothesis is that the species' patchy, fragmented occurrence on discrete sand banks results in detectable genetic separation among populations, which could therefore require individualized management approaches. Alternatively, the species' early life history appears to promote large-scale mixing because *A. dubius* are winter spawners with long embryonic durations and subsequent larval stages that can remain pelagic for up to three months before settling (Suca *et al.*, 2022). Given the strong, predominantly north-to-south surface currents along the NWA, larval transport may indeed suffice to maintain essentially one genetically homogeneous population that could be considered a single entity.

An additional challenge to sand lance research in the NWA is the existence of the congener *Ammodytes americanus* (DeKay, 1842), the American sand lance, which is morphologically and genetically so similar to *A. dubius* that the two species have undoubtedly been confused in the literature (Staudinger *et al.*, 2020). Their meristic differentiation is subtle and incomplete (Nizinski *et al.*, 1990), their mitochondrial genomes indicate shallow divergence (Horne *et al.*, 2016), and standard mtDNA barcoding fails to distinguish them as two species (McCusker *et al.*, 2013). Most research therefore proceeds by assuming complete allopatry, which means that all specimens caught on offshore sand banks are considered *A. dubius*, whereas all specimens encountered in nearshore waters are assumed to be *A. americanus* (Nizinski *et al.*, 1990). If true, this would constitute an example of remarkable niche partitioning, but a clear test of this hypothesis is still lacking.

To examine the potential population structure of *A. dubius* across the NWA and its separation from *A. americanus*, we used low-coverage whole genome sequencing (lcWGS), which has become a popular and powerful approach for exploring patterns of genetic differentiation (Therkildsen and Palumbi, 2017; Lou *et al.*, 2021). A low average per sample sequencing depth of 1–2 × comprises the trade-off that makes analysing whole genomes of hundreds of specimens affordable. While this sequencing depth is too low to reliably call genotypes, analysis of low-coverage data within a probabilistic framework (that explicitly incorporates the uncertainty about individual genotypes in downstream analysis) allows robust inference of many population genetic parameters, including allele frequencies and population differentiation (Buerkle and Gompert, 2013; Fumagalli, 2013; Nevado *et al.*, 2014; Lou *et al.*, 2021). To take advantage of lcWGS, we first assembled a high-quality reference genome for *A. dubius* (the first in the Ammodytidae family) and then used the approach to examine hundreds of collaboratively obtained *A. dubius* specimens ranging from locations in the Mid-Atlantic Bight, Gulf of Maine, the Grand Banks, and Greenland, in addition to outgroup congeners *A. americanus* and the Pacific *A. personatus*. Our analyses revealed for the first time the existence of two genetically distinct *A. dubius* population clusters and suggested unambiguous genetic delineation between the two NWA congeners, which will aid future research and conservation of NWA sand lances.

Material and methods

Sample origin

Sample locations spanned most of the geographical distribution of *A. dubius* in the NWA (Figure 1, Table 1), with specimens obtained from seven different regions referred to here as (1) Mid-Atlantic Bight (MAB), (2) Georges Bank, (3) Cape Cod, (4) Stellwagen Bank, (5) Gulf of St. Lawrence (GSL), (6) Grand Banks, and (7) the Godthåbsfjord (Nuuk, Greenland). All specimens from MAB, Georges Bank, Cape Cod, and Stellwagen Bank were obtained during dedicated research cruises on the *R/V Auk* using a 1.3 × 0.7 m beam trawl (6 mm mesh) towed over ground at three knots for 15 min. All available samples from MAB, Georges Bank, and Stellwagen Bank were sequenced ($n = 72$). From Cape Cod, we randomly selected 45 of the 60 available specimens for sequencing. Specimens from the GSL were collected in 2019 with the *CCGS Teleost* by the Canadian Department of Fisheries and Oceans' (DFO) annual multidisciplinary survey in northern and southern areas. A four-sided Campelen 1800 shrimp trawl (codend = 13 mm) and an Atlantic Western IIA Survey trawl (codend = 19 mm) were used in the northern and southern GSL surveys, respectively. Subsamples of 20 (northern GSL) and 21 (southern GSL) specimens were randomly selected for sequencing. Grand Banks specimens were sampled at 22 sites with the *CCGS Teleost* (Camplelen 1800 shrimp trawl) in October of 2019, and 77 individuals from eight trawl sites were chosen for sequencing. Specimens obtained from Godthåbsfjord were not adults, but larvae of ~10 mm TL, sampled via plankton

Table 1. Sample overview for low-coverage whole genome sequencing, including species, country, region, year, average latitude and longitude, number of sites sampled via trawl, seine, or ichthyoplankton tows, and the final number of sequenced samples (N).

Species	Country	Region	Year	Latitude (N) Longitude (W)	Sites	N
Ammodytes dubius	USA	Cape Cod, Gulf of Maine	2017	41° 56.346’ 69° 59.334’	2	45
		Georges Bank	2017	40° 48.93’ 68° 31.482’	2	10
		Mid-Atlantic Bight	2018	39° 8.742’ 74° 13.224’	1	7
		Stellwagen Bank, Gulf of Maine	2017	42° 27.234’ 70° 20.166’	5	28
			2019	42° 27.234’ 70° 20.166’	1	27
	Canada	Grand Banks	2019	45° 34.08’ 50° 36.42’	22	74
		Gulf of St. Lawrence, North	2019	49° 56.068’ 60° 55.804’	24	20
		Gulf of St. Lawrence, South	2019	47° 39.428’ 63° 20.567’	20	21
	Greenland	Godthåbsfjord	2018	64° 27.132’ 51° 28.032’	1	30
Ammodytes personatus	USA	Salish Sea	2019	48° 42.456’ 122° 42.372’	2	10
Ammodytes americanus	USA	Casco Bay, Gulf of Maine	2017	43° 44.61’ 69° 59.172’	1	14

tows (335 µm mesh) on 14 June 2018. A subsample of 30 larvae was randomly selected for sequencing.

Outgroup samples of two congeners, *A. americanus* and *A. personatus,* were included in the analysis (Table 1). *A. americanus* was collected by beach seine in 2017 in Casco Bay, totalling 15 specimens from one site, which were all sequenced. Pacific *A. personatus* were collected by sediment grab from two nearshore sites in the Salish Sea, and all ten specimens were sequenced. In total, 311 individuals were selected for sequencing. Frozen adult specimens were briefly thawed to obtain caudal fin clips with ethanol and flame-sterilized scissors to avoid cross-contamination. Fin clips were subsequently stored in 95% ethanol. Larvae from Godthåbsfjord were used whole to ensure sufficient DNA yield.

DNA extraction and sequencing

DNA was extracted from all samples using the Omega Bio-Tek E-Z 96 Tissue DNA Kit (Omega Bio-Tek, Inc., Norcross, GA, USA), following the company’s standard tissue protocol sheet. Successful extractions had DNA concentrations exceeding 0.4 ng µl^{-1}, which were measured using a Qubit high sensitivity assay kit (Invitrogen®). Sequencing libraries were prepared by Tn5 transposase coupled to magnetic beads (Tn5 expression as described in Picelli *et al.*, 2014), producing a uniquely barcoded library for each individual (similar to the method used in Therkildsen and Palumbi, 2017). This approach allows for pooling samples for sequencing while retaining individual-level information required for population structure analysis (Lou *et al.*, 2021). Samples were pooled and sequenced across three lanes (paired ends, 150 bp) on an Illumina HiSeq platform at Novogene Inc., CA. We aimed for an average sequencing depth of coverage of 1–2 × for all samples based on a genome size of ∼750 Mb. After assessing data retention on initial sequencing, two additional lanes were sequenced to balance out sequence coverage across individuals.

Assembly of a reference genome

To obtain a high-quality reference genome for *A. dubius*, fin clips of one individual sampled on Stellwagen Bank in November 2020 were flash frozen in liquid nitrogen and stored at −80°C. The sample was shipped to Dovetail Genomics (CA) for sequencing, assembly, and scaffolding using their standard pipeline that involves PacBio™ HiFi™ long reads to assemble contigs followed by scaffolding using proprietary Omni-C proximity ligation technology and HiRise scaffolding software (ESM1 contains further details).

Data filtering

Raw fastq files were first de-multiplexed (reads assigned to individuals) using bbmap v37.41 (https://github.com/BioInfoTools/BBMap) with the demuxbyname.sh script. We clipped off adapter sequences with Trimmomatic v0.39 (Bolger *et al.*, 2014) and used FastQC v0.11.5 (Andrews, 2010) and MultiQC v1.10.1 (Ewels *et al.*, 2016) to check the sequencing quality across all individuals. Because the average phred-scaled base quality score was >20 even at the ends of the reads in almost all samples and because we analysed the sequence data in a probabilistic framework that takes base quality scores into account (thus weigh bases with poor quality scores lower), we did not quality trim ends of reads, but instead ignored low-quality bases in our downstream analysis. All reads were mapped to the reference genome using Bowtie2 v2.3.5.1 (Langmead and Salzberg, 2012) with the “very-sensitive” preset and then filtered using Samtools v1.9 (http://www.htslib.org/) to remove reads with a mapping quality <20. Overlapping ends of read pairs were clipped using BamUtil v1.0.7 (Breese and Liu, 2013), and duplicated reads were removed using Picard v2.2.1 (http://broadinstitute.github.io/picard/). For samples that received two rounds of sequencing to balance out coverage (97% of individuals), all reads from the same sample were merged into single files for further analysis. After the bioinformatic pipeline, 286 individuals remained for analysis.

SNP calling and population structure analysis

Detection of single nucleotide polymorphisms (i.e. SNP calling) across all individuals was performed using ANGSD v0.933–102 (Korneliussen *et al.*, 2014). To exclude genomic regions with poor read mapping rates and repetitive multicopy elements, we examined the distribution of total read depth per site (across all individuals) and then excluded SNPs $<0.5\times$ and $>2\times$ of the average read depth by setting the minimum and maximum depth filters to 93 (-minDepth93) and 832 (-maxDepth832). Furthermore, the minimum individual filter was set to 77 (-minInd77, or approximately a quarter of the final sample list) to ensure that SNP calling at a given site had a sufficient number of samples. Low-quality base calls and poorly mapped reads were excluded (-minQ 20 and -minMapQ 20). If a given site had a minor allele frequency significantly different from 0, it was deemed polymorphic. We kept SNPs with a minor allele frequency of 1% using a minimum minor frequency filter (-minMaf 0.01) with a *p*-value filter of 10^{-6} for polymorphic sites (–SNP_pval 1e-6). Genotype likelihoods were also generated during the SNP calling step (–GL 1; –doGlf 2). The beagle-formatted genotype likelihoods were then processed with PCAngsd v.099 using the PCA setting to create an individual-level covariance matrix. The final, formatted output file from PCAngsd was then imported into R (R Development Core Team, 2018) for eigendecomposition, extracting two principal components that explained a cumulative 22.1% of the total variability.

In addition to this analysis of variation across nuclear SNPs, we also examined patterns of sequence variation across the mitochondrial genome. Because most cells have far more copies of the mitochondrial genome than the nuclear genome (and mitochondrial genomes are haploid), shotgun sequencing typically results in 50–200 × higher read depth for the mitochondrial genome than the nuclear genome, and our low-coverage (of the nuclear genome) sequencing thus allowed us to recover a high-confidence full mitochondrial genome sequence for each individual (Therkildsen and Palumbi, 2017). To obtain these, we mapped reads from each of the 286 individuals to the *A. dubius* mitochondrial reference genome (NC_028 529, NCBI) from Horne *et al.* (2016), followed by using ANGSD (-doCounts 1) for each individual to count observed alleles at each position of the mitochondrial genome. We then used R to generate a consensus sequence for each individual, with the major allele chosen to be the consensus allele for each position if the sequencing depth was $\geq 4\times$ and if the major allele frequency was ≥ 0.75 (otherwise the position was given an N). We then produced minimum-spanning haplotype networks from these consensus sequences using PopART (Bandelt *et al.*, 1999; Leigh and Bryant, 2015).

Next, we removed all 29 samples from the two congener outgroups (*A. americanus* and *A. personatus*) from the sample list and re-ran the nuclear SNP calling using ANGSD with only *A. dubius* samples and a more stringent minimum minor frequency filter of 5% (-minMaf 0.05), a minimum depth filter of 350 (-minDepth 350), and a maximum depth of 604 (-maxDepth 604). Linkage disequilibrium was calculated using ngsLD v1.1.1 (Fox *et al.*, 2019). After visualizing LD decay, we chose a filter of 5 kbp to generate a LD-pruned data set using the prune_graph.pl function. To explore the LD pattern in areas of high divergence, we ran ngsLD with—max_kb_dist disabled and visualized results using the LDheatmaps package in R. To determine the degree of genetic differentiation among geographic regions, we first conducted population-level F_{ST} analyses using the genotype likelihood data from ANGSD. For these pairwise comparisons, we first estimated sample allele frequency likelihoods using the -doSaf 1 option in ANGSD, which were then used to estimate F_{ST} values at each SNP for each pair of geographic regions with the realSFS F_{ST} function. To examine how levels of genetic differentiation vary across the genome, we used R to generate Manhattan plots of F_{ST} between the northern and southern regions. To test the statistical significance of the correlation between the genetic and geographic distances, we performed Mantel tests in R (Mantel.rtest). The genetic distance (GD) was calculated for each region from genome-wide average F_{ST} estimates as

$$GD = \frac{F_{ST}}{(1 - F_{ST})}.$$

The geographic distance matrix was calculated as a least-cost path using a publicly available bathymetry grid (General Bathymetric Charts of the Ocean, https://download.gebco.net). Using Marmap (Pante and Simon-Bouhet, 2013), we read bathymetry data and set the minimum depth for connecting paths to 1 m. Genetic and geographic distance matrices were input into the Mantel.rtest function in the ade4 package (Dray and Dufour, 2007) in R to conduct an isolation-by-distance analysis (Mantel test).

Results

A total of 6339586 PacBio reads (amounting to 110.9 Gb of data and estimated 139 × genome coverage) were used for the assembly of the *A. dubius* reference genome. After filtering and scaffolding with Omni-C data, the final Dovetail HiRise Assembly contained 748559117 bp assembled into chromosome-level contigs (largest scaffold = 35270950 bp, N90 = 22220510 bp) with L50 and L90 values of 12 and 22, respectively, and the Omni-C contact map indicated that *A. dubius* has 24 chromosomes. The Benchmarking Universal Single-Copy Orthologs (BUSCO) score was 92.16% based on 255 total BUSCO groups searched (Table 2, Supplementary Material ESM1).

Our low-coverage, whole genome sequencing of 286 individuals yielded a median (10%; 90%) raw base count of 2.02 Gb per individual (1.81 Gb; 2.75 Gb), amounting to an average raw read depth per individual of 2.94× (Supplementary Figure S1). After mapping to the reference genome and deduplication, the median (10%; 90% percentile) retained data amounted to a base count of 0.97 Gb per individual (0.93 Gb; 1.23 Gb), which is equivalent to a final average read depth per individual of 1.40×. Data yields were similar among *A. dubius* specimens from the different geographic regions and comparable to *A. americanus*, but notably higher and more variable for *A. personatus* (Supplementary Figure S1). For the mitochondrial genome, the average read depth per individual was 107× (range: 20–470×).

The genome-wide PCA based on 38738134 SNPs revealed strong genetic differentiation between congeners *A. dubius*, *A. americanus*, and *A. personatus* in our samples (Figure 2a). Along PC1 (explaining 18.4% of the total variation), *A. dubius* separated from *A. americanus*/*A. personatus*, while all three congeners were well separated from each other along PC2 (3.7%). Genome-wide average F_{ST} values between pairs of congener clusters were high, with $F_{ST} = 0.28$ between *A. dubius* and *A. americanus*, $F_{ST} = 0.54$ between *A. dubius* and

Table 2. Key statistics of the *A. dubius* reference genome assembly and scaffolding results (Dovetail Genomics).

Statistic	Primary (WTDBG2)	Primary filtered = Input assembly	HiRise assembly
Total length (bp)	831 218 799	748 472 917	748 559 117
N50 (bp)	2 975 273	3 578 647	31 318 865
N90 (bp)	110 727	394 186	22 220 510
L50	65	52	12
L90	656	293	22
Largest scaffold (bp)		18 355 163	35 270 950
Number of scaffolds		1 697	835
Number of scaffolds > 1kbp		1 688	826
Number of N's per 100 kbp		0.02	11.54
Complete single copy BUSCOs		235 (92.2%)	235 (92.2%)

Primary (WTDBG2) is the input assembly, which was then filtered and used as input for producing a HiRise assembly (the final assembly we used as the reference genome). N50, N90: sequence length of the shortest contig at 50–90% of the total assembly length, respectively. L50, L90: the smallest number of contigs whose length sum made up 50–90% of the genome size, respectively. Benchmarking Universal Single-Copy Orthologs (BUSCO) scores were computed based on 255 total BUSCO groups searched (see Supplementary Material ESM1 for further details).

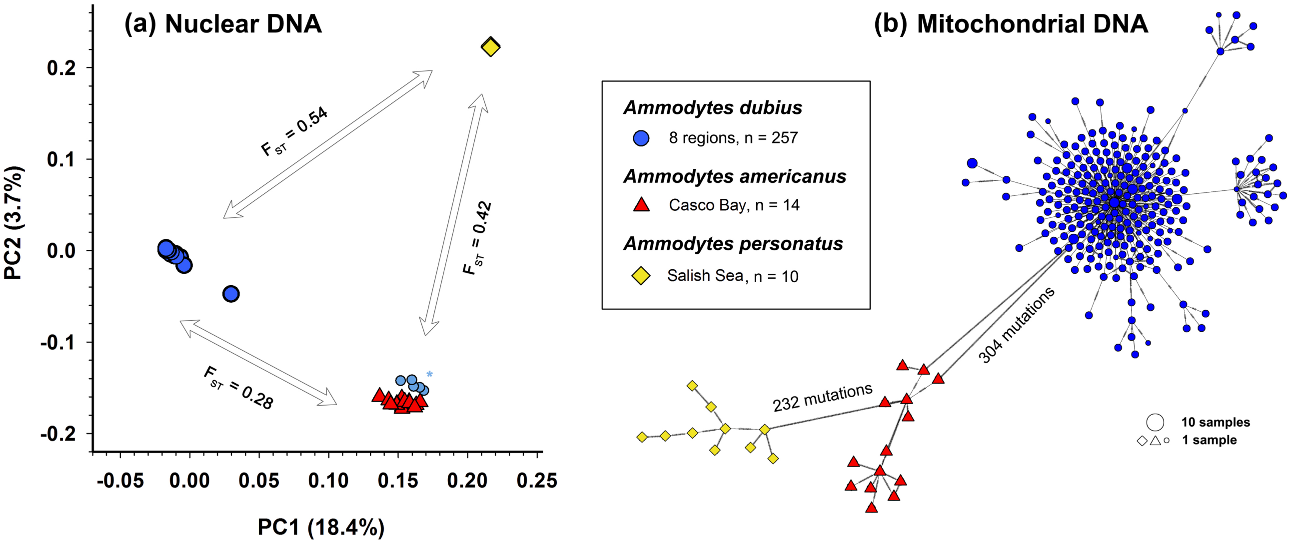

Figure 2. Genetic separation between the three *Ammodytes* congeners is based on (a) nuclear DNA and (b) mitochondrial DNA. (a) Principal component analysis of the SNP-based, individual-level covariance matrices with superimposed F_{ST} values for each interspecies comparison. The asterisk denotes five specimens from the northern GSL that were caught as *A. dubius* but clearly grouped with *A. americanus* and were therefore likely misidentified and excluded from subsequent analyses. (b) Mitochondrial haplotype network of all analysed specimens, with the number of mutations separating congeners.

A. personatus, and $F_{ST} = 0.42$ between *A. americanus* and *A. personatus* (Figure 2a). Five specimens sampled as "*A. dubius*" in the northern GSL tightly clustered with *A. americanus* (Figure 2a); these specimens were likely "misidentified *A. dubius*" and thus removed from further analysis. The separation of the three congeners was also evident in their mitochondrial genome (~16 kbp, Figure 2b), with the resulting haplotype network map showing *A. dubius* separated from *A. americanus* by 304 mutations and *A. americanus* separated from *A. personatus* by 232 mutations. Within *A. dubius*, haplotypes showed no geographical pattern.

However, considering only *A. dubius* nuclear genomic data ($n = 257$), a second PCA based on 6732117 SNPs revealed intra-specific clustering along PC1 (explaining 1.8% of the total variation), because all specimens from southern geographic regions (MAB, Georges Bank, Cape Cod, Stellwagen Bank) formed one distinct cluster, whereas all specimens from northern geographic regions (Grand Banks, GSL, Godthåbsfjord) formed another distinct cluster (Figure 3a). The southern and northern population clusters showed no overlap along PC1. The pairwise F_{ST} analyses showed the highest genetic differentiation between MAB and all other geographic regions ($F_{ST} = 0.030$–0.046), although this may partially be due to lower sample numbers for MAB ($N = 7$; Figure 3b). Within the southern cluster, specimens from Stellwagen Bank and Cape Cod were genetically similar ($F_{ST} = 0.007$), whereas both locations showed greater genetic differentiation to Georges Bank ($F_{ST} = 0.017$–0.019, Figure 3b). Within the northern cluster, pairwise F_{ST} values indicated little differentiation and no geographic pattern, with the least differentiation between northern GSL and Grand Banks ($F_{ST} = 0.003$) and the most differentiation between southern GSL and Godthåbsfjord ($F_{ST} = 0.013$, Figure 3b).

Our isolation-by-distance analysis for *A. dubius* found no significant correlation ($p = 0.19$) between genetic and geographic distance (Mantel tests) when we tested across

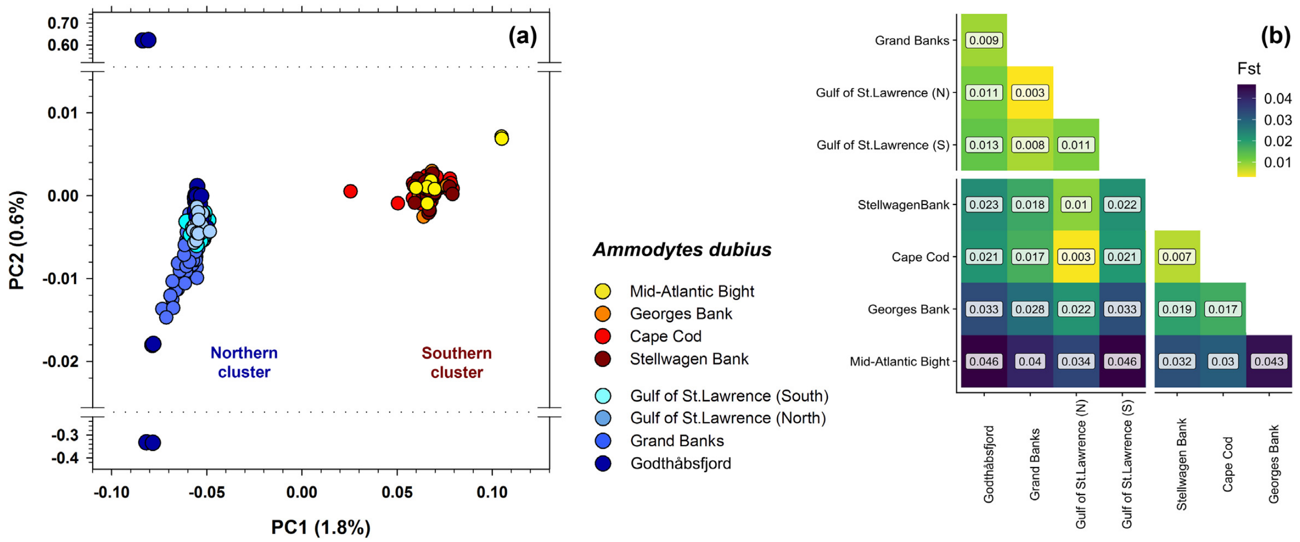

Figure 3. Intra-specific *A. dubius* genomic population structure (excluding congeners). (a) PCA including 257 specimens, showing complete separation of northern and southern population clusters along PC1 (notice the breaks in the PC2-axis to visualize four outliers from Godthåbsfjord, Greenland). (b) Pairwise region-level F_{ST} comparisons for *A. dubius* from all eight NWA regions.

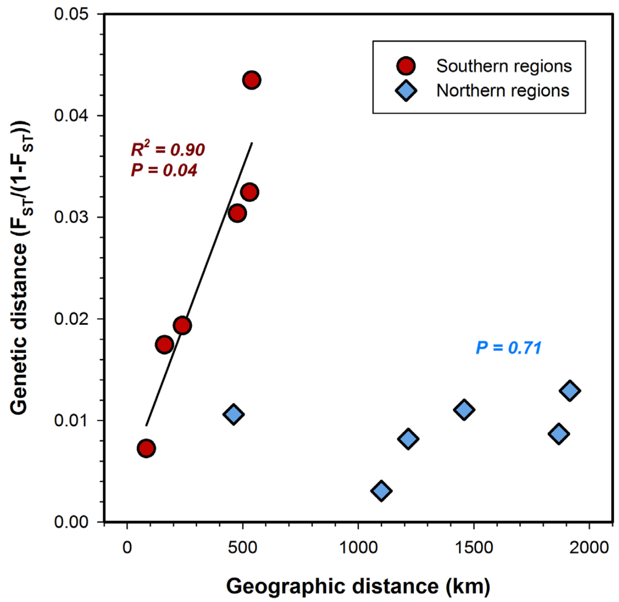

Figure 4. Relationship between GD (calculated from F_{ST} values) and geographic distance for all *A. dubius* sampling groups. Red circles and blue diamonds represent groups within the southern and northern population clusters, respectively. Mantel tests were conducted for each cluster separately, with test statistics shown within the plot.

all populations. However, when the northern and southern clusters were examined separately, we found that GD increased with geographic distance within the southern cluster ($p = 0.04$, $r^2 = 0.90$; Figure 4), but not within the northern cluster, despite covering much larger geographical distances ($p = 0.71$; Figure 4).

When we examined genome-wide differentiation patterns within northern and southern population clusters, we saw uniformly low $F_{ST} < 0.02$ values for all within-cluster population comparisons (Supplementary Figures S1–S2). However, when comparing northern and southern population clusters (Figure 5a), we discovered narrow peaks of highly elevated differentiation ($F_{ST} > 0.5$) on many different chromosomes and a large contiguous block of highly elevated differentiation in strong LD spanning more than two-thirds of chromosome 21 and a smaller block of elevated differentiation on chromosome 24 (Figure 5b and c). While most chromosomes had an average F_{ST} value of 0.002, chromosomes 21 and 24 had average F_{ST} values of 0.032 (16 × average) and 0.009 (4.5 × average), respectively. We therefore conducted additional PCAs including only SNPs from chromosomes 21 or 24 (Figure 6a and b), another PCA including only SNPs from all other chromosomes (Figure 6c), and a final PCA using an n-equalized ($n \leq 30$), LD-pruned SNP dataset (Figure 6d). This confirmed that the strong separation of *A. dubius* into northern and southern population clusters along PC1 was largely but not solely due to the genomic differentiation concentrated on chromosomes 21 and 24, because the clusters remained detectable across the other chromosomes and in the LD-pruned dataset.

Discussion

The rapid rise of low-coverage whole genome sequencing has opened up vast new possibilities for studying non-model organisms across the tree of life (Lou *et al.*, 2021; e.g. Akopyan *et al.*, 2022; Bahbahani and Almathen, 2022). Here, we used the approach to reveal the population structure of Northern sand lance (*A. dubius*), a critical but understudied NWA forage fish with intriguing behaviour and likely high climate vulnerability (Staudinger *et al.*, 2020; Suca *et al.*, 2021; Baumann *et al.*, 2022). By analysing variation across millions of SNPs across hundreds of *A. dubius* genomes spanning the species' geographical range, we discovered the existence of a large northern and smaller southern population cluster while also suggesting a clear genetic delineation between offshore *A. dubius* and inshore *A. americanus*.

A. dubius population structure

Our SNP-based PCA suggested that *A. dubius* from Cape Cod, Stellwagen Bank, Georges Bank, and the Mid-Atlantic Bight form one genetic cluster that is distinct from conspecifics fur-

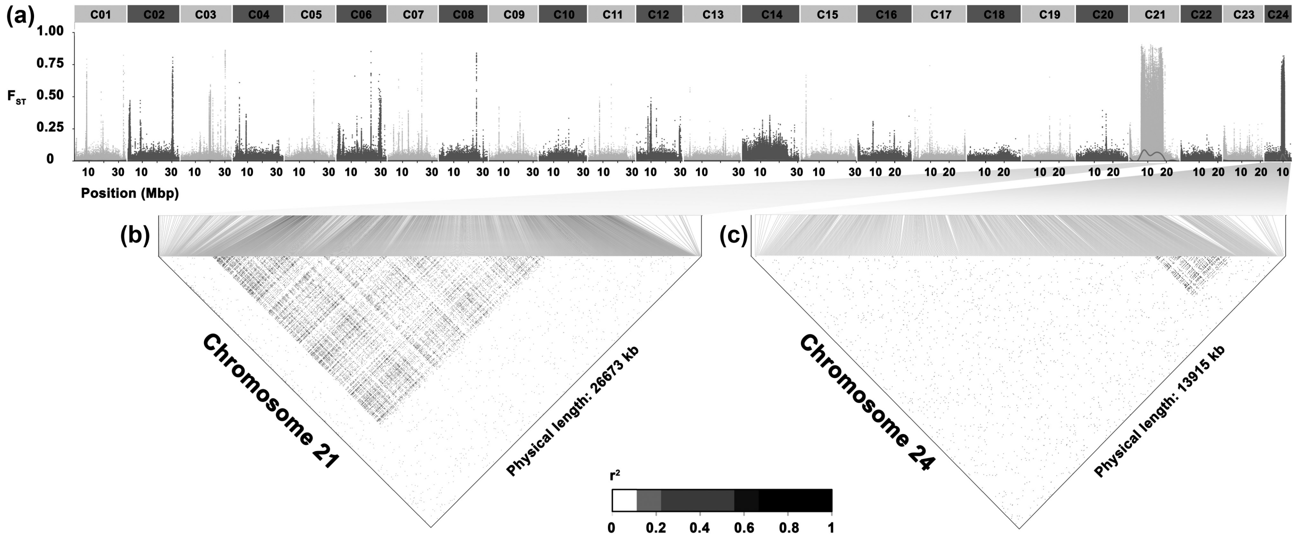

Figure 5. (a) Manhattan plot showing how levels of genetic differentiation (F_{ST}) between northern and southern population clusters vary across the genome. Points denote individual SNPs, organized by position in the genome, with a subsetted panel for each chromosome. (b) Pairwise SNP LD heatmap for chromosome 21. (c) Pairwise SNP LD heatmap for chromosome 24. For (b) and (c) a 0.5% subset of SNPs was used.

ther north in the Gulf of St. Lawrence, Grand Banks, and the west coast of Greenland (Figure 1). Hence, the apparent genetic break between the two clusters likely occurs somewhere on the Scotian Shelf, which is strikingly consistent with previously documented genetic breaks in other NWA species. For example, van Wyngaarden *et al.* (2017) showed that sea scallops (*Placopecten magellanicus*) transition between northern and southern genotypes on the Scotian Shelf, and in addition, Lehnert *et al.* (2019) found finer-scale inshore versus offshore genetic clustering. Another comprehensive study looked at scallops, three species of commercially important crustaceans (*Homarus americanus, Carcinus maenas,* and *Pandalus borealis*), and one fish (*Gadus morhua*), which again produced evidence for a cross-taxa genetic break on the Scotian Shelf, despite obvious interspecific differences in life histories, larval durations, or adult mobility (Stanley *et al.*, 2018).

What maintains the Scotian Shelf genetic break in sand lance and other species is a fascinating question. There are no obvious physical or hydrodynamic barriers to limit dispersion there. In fact, the break occurs despite prevailing north-to-south surface transport by the Labrador Current and its southward extensions into the Gulf of Maine and beyond. It appears that most species' planktonic larvae would get readily advected across this genetic transition zone (Oliver *et al.*, 2012), and in the case of *A. dubius*, this seems all but guaranteed. Sand lance are winter spawners; their benthic eggs take several weeks to hatch (Smigielski *et al.*, 1984; Baumann *et al.*, 2022), and larvae remain pelagic for up to three months (Suca *et al.*, 2022). Such extended drift periods allow for long-distance advection and homogeneous mixing, particularly in the winter when NWA currents are usually strongest (Bundy *et al.*, 2014, Staudinger *et al.*, 2020). Adult movements, on the other hand, appear too limited to contribute to sand lance population connectivity. Observations from Stellwagen Bank suggested that, once settled, adult *A. dubius* remain stationary to the bank despite their pelagic feeding. This is consistent with the lesser sandeel, *A. marinus*, in the North Sea, which is deemed non-migratory in most studies (Furness and Tasker, 2000; Arnott and Ruxton, 2002). Therefore, weak to non-existent genetic structures within the North Sea (Jiménez-Mena *et al.*, 2020) or in Norwegian waters (Nævdal and Thorkildsen, 2002) are likely due to genetic connectivity during the pelagic larval stage. In *A. dubius*, the potential for larval mixing is similarly large, specifically via larval drift from north to south throughout the NWA—and yet, southern regions are genetically distinct from northern regions.

Most evidence to date points to temperature as the main driver of the Scotian Shelf genetic break. As Stanley *et al.* (2018) hypothesized, when cold-adapted individuals from northern shelf areas are transported into warmer southern areas, they suffer from higher mortalities compared to their more warm-adapted conspecifics. Lehnert *et al.* (2019) argued that temperature-mediated survival also explains the more fine-scale genetic structuring between sea scallops in shallow (warm) and deep (cold) shelf areas. In green crabs, Coyle *et al.* (2019) present evidence that northern and southern populations have divergent temperature adaptations that map to unique mitochondrial haplotypes. For *A. dubius*, we suggest that winter sea surface temperature (i.e. the average seasonal SST minimum; Assis *et al.*, 2018) is a potentially predictive thermal break, because it clearly delineated northern populations (≤3°C) from southern populations (≥5°C, Figure 1, Supplementary Figure S4). We therefore hypothesize that it is not current-mediated mixing but the temperature-mediated survival of progeny that maintains the genetic break in *A. dubius*. Differential mortality of northern and southern genotypes limits the "realized" gene flow across the thermal gradient of the Scotian Shelf, thus causing the observed genetic break.

While we have yet to learn what genes actually distinguish northern and southern sand lance populations, our study took a first step by examining whether genetic differentiation is diffuse or localized to specific regions of the genome. Anchoring our analysis onto the new reference genome of *A. dubius* allowed examining how loci-specific pairwise F_{ST} comparisons between northern and southern population vary along the genome. This revealed that genetic differentiation is highly elevated in large blocks on chromosomes 21 and 24. These

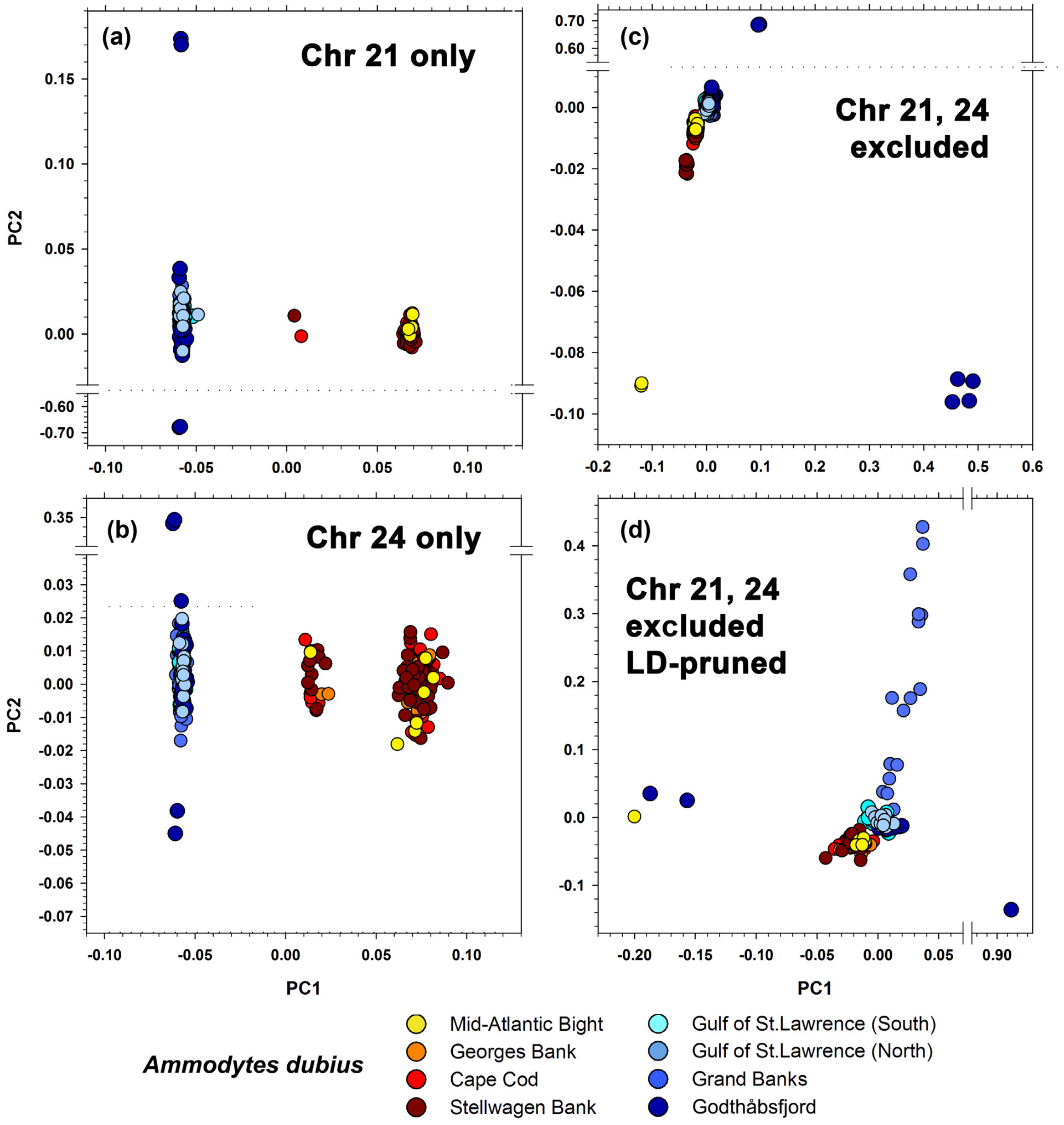

Figure 6. Intra-specific *A. dubius* genomic population structure based on principal component analysis of the SNP-based, individual-level covariance matrices. (a) PCA including only SNPs contained on chromosome 21. (b) PCA including only SNPs contained on chromosome 24. (c) PCA including only SNPs not contained on chromosomes 21 or 24. (d) PCA excluding chromosomes 21 and 24 and using n-equalized, LD-pruned data (notice the breaks in the PC2-axis to visualize outliers from Godthåbsfjord, Greenland).

chromosomes will be targeted in future investigations because they likely contain the important alleles that determine sand lance fitness under contrasting thermal regimes. Highly localized areas of differentiation in the genome, often structurally caused by inversions, are known to play a major role in local adaptation in the ocean despite gene flow, as demonstrated recently in another NWA forage fish (*Menidia menidia*; Wilder *et al.*, 2020; Akopyan *et al.*, 2022) and other marine species such as Atlantic cod (*Gadus morhua*; Puncher *et al.*, 2019), periwinkle snails (*Littorina saxatilis*; Koch *et al.*, 2021), and seaweed flies (*Coelopa frigida*; Mérot *et al.*, 2021). Localized genetic differentiation also comprises yet another argument in favour of whole-genome over targeted sequencing approaches, given the likelihood of missing important population structure when analysing only a limited, random number of loci (Jiménez-Mena *et al.*, 2020; Lou *et al.*, 2021).

Because most sand lance species display a peculiar burying behaviour that effectively constrains them to discontinuous, coarse-grain sand banks, one might expect genetic and geographic distances to be highly correlated in *A. dubius*. Indeed, within the smaller southern population cluster, GDs significantly increased with geographic distances (min/mean/max distance = 83/338/539 km), indicating that currents cannot completely homogenize sand lance progeny from different banks. For example, Georges Bank has a semi-permanent, retentive eddy that clearly explained sand lance drift patterns in a modelling study (Suca *et al.*, 2022). This is consistent with the observable genomic separation seen between Georges Bank and the Gulf of Maine sand lance and consistent with evidence from sea scallops (Tremblay *et al*., 1994). For the much larger, northern population cluster, the corresponding Mantel tests clearly indicated genetic homogeneity despite the vast geographic distances between sampling sites (min/mean/max distance = 460/1336/1916 km). This probably indicates effective mixing of progeny, aided by strong surface currents and likely longer larval drift durations in colder waters. However, detailed demographic modelling is needed to rule out the possibility that gene flow is also spatially restricted in the northern cluster and the genetic homogeneity is retained because of minimal drift in large population sizes coupled with relatively recent postglacial recolonization.

Separation of *A. dubius* and *A. americanus*

A secondary objective of our study was to evaluate whether lcWGS could distinguish the two NWA congeners *A. dubius* and *A. americanus,* which have only been recognized as separate species since the 1980s (Richards, 1982; Winters and Dalley, 1988). The species' overlapping latitudinal ranges and meristic characters hinder reliable identification (Winters and Dalley, 1988), and previous genetic approaches have had limited success. For example, McCusker *et al.* (2013) focused on the mitochondrial cytochrome c oxidase 1 (COI) barcoding region but concluded that the method did not separate the two *Ammodytes* species, possibly because they split too recently. Horne *et al.* (2016) sequenced two complete mitochondrial genomes, noting very limited sequence divergence between the two species. In this study, both mitochondrial and nuclear genomes clearly differentiated *A. dubius* from *A. americanus*, the latter yielding pairwise F_{ST} values (mean $F_{ST} = 0.182$) that were about one order of magnitude greater than those between *A. dubius* populations (mean $F_{ST} = 0.025$). The genetic differentiation occurs despite close geographic proximity, as the nearshore sampling site for *A. americanus* (Casco Bay, ME, USA) is only 143 km apart from the offshore Stellwagen Bank.

We draw two important conclusions from these observations. First, we found that every sequenced sand lance specimen caught on Stellwagen Bank was indeed *A. dubius*, which lends confidence to past and future research efforts there (Murray *et al*., 2019; Silva *et al*., 2020; Suca *et al*., 2021). Second, at one site close to shore in the Strait of Belle Isle (northern GSL), we actually detected five specimens that were caught as "*A. dubius*", but their genome clearly clustered with *A. americanus*. We therefore conclude that the remarkable spatial niche partitioning into inshore *A. americanus* and offshore *A. dubius* sand lance species is a valid assumption, but if the two habitats are poorly separated, mixing and misidentifications can occur. Genomic tools could be used for developing targeted assays that would help management and conservation efforts of both species, including better resolving their spatial overlap and potential for hybridization. The latter has not been documented yet, perhaps because suitable nearshore and offshore habitats for sand lance are mostly well separated by deeper waters and/or unsuitable substrates (e.g. mud, rock) where sand lance cannot bury and thus do not occur, hence largely preventing hybridization zones.

In summary, our study contributed valuable basic knowledge about the population structure of northern sand lance on the NWA, showing that the species' long larval duration and the shelf's strong surface currents indeed achieve genetic homogenization across vast spatial scales (10^3 km) within the northern population cluster. On the other hand, we clearly detected the Scotian Shelf genetic break in our samples and—consistent with previous research—suggest that it reflects temperature-mediated survival differences in progeny (realized connectivity). Northern sand lance face an uncertain future due to their direct and indirect sensitivity to climate change and threats to their habitat (Silva *et al*., 2020; Staudinger *et al*., 2020; Suca *et al*., 2021; Baumann *et al*., 2022); hence, northern and southern *A. dubius* genotypes likely require separate, ecosystem-based management strategies to protect their crucial role in temperate to polar NWA ecosystems.

Acknowledgements

We are truly grateful to all who facilitated sand lance sampling for this study, including the Department of Fisheries and Oceans, Canada, the crews of the *RV Auk* and *CCGS Teleost*, as well as T Silva, M Thompson, and P Hong. We thank V de Carufel and F Heinrich for help with sampling in Greenland, P Warner and J Dunn for field sampling in Casco Bay, ME, as well as NT Schwarck, the Washington Ocean Acidification Center, and the Washington Department of Fish and Wildlife for help with Pacific sand lance sampling. We are grateful to H Borchardt-Wier for her assistance with DNA extraction and library preparation, and to the three anonymous reviewers for their helpful comments and suggestions.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Funding

This work was supported by a Bureau of Ocean Energy Management (BOEM) grant to D. Wiley, H. Baumann, J. Llopiz, and L. Kaufman under IA agreement M17PG0019. LJ was also partially funded by NSF-OCE 1756316 to N.O.T. and H.B. M. K. was supported by ERC grants 639096, "HybridMiX", and 101069216, "HAPLOTAGGING". The *A. dubius* reference genome was assembled by Dovetail Genomics under a "*Tree of Life*" grant to H.B. and N.O.T., with additional contributions from H.B., N.O.T., the UConn College of Liberal Arts and Sciences, and the Department of Marine Sciences.

Author contributions

L. Jones, N. Lou, H. Baumann, and N. Therkildsen devised the study, analysed the data, and wrote the manuscript. C. Murray, D. Robert, C. Bourne, C. Bouchard, D. Carlon, and D. Wiley contributed specimen samples. F. Chan and M. Kučka provided novel reagents for library preparation. All authors edited and approved the final version of the manuscript.

Data availability statement

The *A. dubius* reference genome, metadata, and all individual sequence read archives (SRA, $n = 315$, fastq.gz archives) are publicly available from the National Center for Biotechnology Information (NCBI) under the BioProject PRJNA834854 (=Accession) https://www.ncbi.nlm.nih.gov/bioproject/PRJNA834854.

References

Akopyan, M., Tigano, A., Jacobs, A., Wilder, A. P., Baumann, H., and Therkildsen, N. O. 2022. Comparative linkage mapping uncovers recombination suppression across massive chromosomal inversions associated with local adaptation in Atlantic silversides. Molecular Ecology, 31: 3323–3341.

Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Babraham Bioinformatics, Babraham Institute, Cambridge.

Arnott, S. A., and Ruxton, G. D. 2002. Sandeel recruitment in the North Sea: demographic, climatic and trophic effects. Marine Ecology Progress Series, 238: 199–210.

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., and De Clerck, O. 2018. Bio-oracle v2. 0: extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography, 27: 277–284.

Bahbahani, H., and Almathen, F. 2022. Homogeneity of Arabian peninsula dromedary camel populations with signals of geographic distinction based on whole genome sequence data. Scientific Reports, 12: 1–11.

Bandelt, H. J., Forster, P., and Röhl, A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16: 37–48.

Baumann, H., Jones, L. F., Murray, C. S., Siedlecki, S., Alexander, M., and Cross, E. L. 2022. Impaired hatching exacerbates the high CO_2 sensitivity of embryonic sand lance, *Ammodytes dubius*. Marine Ecology Progress Series, 687: 147–162.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics, 30: 2114–2120.

Breese, M. R., and Liu, Y. 2013. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics, 29: 494–496.

Buerkle, C. A., and Gompert, Z. 2013. Population genomics based on low coverage sequencing: how low should we go? Molecular Ecology, 22: 3028–3035.

Bundy, A, Themelise, D, Sperl, J, and den Heyer, N. 2014. Inshore Scotian Shelf ecosystem overview report: status and trends. DFO Canadian Science Advisory Secretariat Research Document, 65: 225.

Coyle, A. F., Voss, E. R., Tepolt, C. K., and Carlon, D. B. 2019. Mitochondrial genotype influences the response to cold stress in the European green crab, *Carcinus maenas*. Journal of Experimental Biology, 222: jeb203521.

Daunt, F., Wanless, S., Greenstreet, S. P., Jensen, H., Hamer, K. C., and Harris, M. P. 2008. The impact of the sandeel fishery closure on seabird food consumption, distribution, and productivity in the northwestern North Sea. Canadian Journal of Fisheries and Aquatic Sciences, 65: 362–381.

Development Core Team, R. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Dray, S., and Dufour, A.-B. 2007. The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22: 1–20.

Engelhard, G. H., Peck, M. A., Rindorf, A. C., Smout, S., van Deurs, M., Raab, K., Andersen, K. H. *et al.*. 2014. Forage fish, their fisheries, and their predators: who drives whom? ICES Journal of Marine Science, 71: 90–104.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 32: 3047–3048.

Fox, E. A., Wright, A. E., Fumagalli, M., and Vieira, F. G. 2019. ngsLD: evaluating linkage disequilibrium using genotype likelihoods. Bioinformatics, 35: 3855–3856.

Fumagalli, M. 2013. Assessing the effect of sequencing depth and sample size in population genetics inferences. PLoS One, 8: e79667.

Furness, R. W. 1999. Does harvesting a million metric tons of sand lance per year from the North Sea threaten seabird populations. Ecosystem Approaches for Fisheries Management, 16: 407–424.

Furness, R. W., and Tasker, M. L. 2000. Seabird-fishery interactions: quantifying the sensitivity of seabirds to reductions in sandeel abundance, and identification of key areas for sensitive seabirds in the North Sea. Marine Ecology Progress Series, 202: 253–264.

Horne, J. B., Mcbride, M. C., Lighten, J., Bradbury, I. R., and Bentzen, P. 2016. Annotated mitochondrial genome assemblies for two sand lances (genus: *Ammodytes*) from the northwest Atlantic. Mitochondrial DNA Part A, 27: 4607–4608.

ICES. 2010. Report of the Benchmark Workshop on Sandeel (WKSAN). ICES Document CM 2010/ACOM:57. 201 pp.

Jiménez-Mena, B., Le Moan, A., Christensen, A., van Deurs, M., Mosegaard, H., Hemmer-Hansen, J., and Bekkevold, D. 2020. Weak genetic structure despite strong genomic signal in lesser sandeel in the North Sea. Evolutionary Applications, 13: 376–387.

Koch, E. L., Morales, H. E., Larsson, J., Westram, A. M., Faria, R., Lemmon, A. R., Lemmon, E. M. *et al.*. 2021. Genetic variation for adaptive traits is associated with polymorphic inversions in *Littorina saxatilis*. Evolution letters, 5: 196–213.

Korneliussen, T. S., Albrechtsen, A., and Nielsen, R. 2014. ANGSD: analysis of next generation sequencing data. Bmc Bioinformatics [Electronic Resource], 15: 356.

Langmead, B., and Salzberg, S. L. 2012. Fast gapped-read alignment with bowtie 2. Nature Methods, 9: 357–359.

Lehnert, S. J., DiBacco, C., Van Wyngaarden, M., Jeffery, N. W., Ben Lowen, J., Sylvester, E. V., Wringe, B. F. *et al.*. 2019. Fine-scale temperature-associated genetic structure between inshore and offshore populations of sea scallop (*Placopecten magellanicus*). Heredity, 122: 69–80.

Leigh, J. W., and Bryant, D. 2015. POPART: full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6: 1110–1116.

Lou, R. N., Jacobs, A., Wilder, A. P., and Therkildsen, N. O. 2021. A beginner's guide to low-coverage whole genome sequencing for population genomics. Molecular Ecology, 30: 5966–5993.

Lynam, C. P., Halliday, N. C., Höffle, H., Wright, P. J., van Damme, C. J. G., Edwards, M., and Pitois, S. G. 2013. Spatial patterns and trends in abundance of larval sandeels in the North Sea: 1950–2005. ICES Journal of Marine Science, 70: 540–553.

McCusker, M., Denti, D., Van Guelpen, L., Kenchington, E., and Bentzen, P. 2013. Barcoding Atlantic Canada's commonly encountered marine fishes. Molecular Ecology Resources, 13: 177–188.

Mérot, C., Berdan, E. L., Cayuela, H., Djambazian, H., Ferchaud, A.-L., Laporte, M., Normandeau, E. *et al.*. 2021. Locally adaptive inversions modulate genetic variation at different geographic scales in a seaweed fly. Molecular Biology and Evolution, 38: 3953–3971.

Morrison, S. 2021. Sand lance (*Ammodytes* spp) on the Newfoundland shelf: habitat selection, diel behaviour, and synchrony of dynamics with other forage fish. *In* Department of Biological Sciences, 153pp. University of Manitoba, Central Winnipeg.

Murray, C. S., Wiley, D., and Baumann, H. 2019. High sensitivity of a keystone forage fish to elevated CO_2 and temperature. Conservation Physiology, 7: 1–12.

Nævdal, G., and Thorkildsen, S. 2002. Genetic studies on species composition and population structure of sand eels (genera: *Ammodytes*, *Hyperoplus* and *Gymnammodytes*) in Norwegian waters. Journal of Applied Ichthyology, 18: 124–126.

Nevado, B., Ramos-Onsins, S. E., and Perez-Enciso, M. 2014. Resequencing studies of nonmodel organisms using closely related reference genomes: optimal experimental designs and bioinformatics approaches for population genomics. Molecular Ecology, 23: 1764–1779.

Nizinski, M., Collette, B. B., and Washington, B. B. 1990. Separation of two species of sand lances, *Ammodytes americanus* and *A . dubius*, in the western North Atlantic. Fishery Bulletin, 88: 241–255.

Oliver, ECJ, Sheng, J, Thompson, KR, and Blanco, JRU. 2012. Extreme surface and near-bottom currents in the northwest Atlantic. Natural Hazards, 64: 1425–1446.

Pante, E., and Simon-Bouhet, B. 2013. marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. PLoS One, 8: e73051.

Picelli, S., Björklund, Å. K., Reinius, B., Sagasser, S., Winberg, G., and Sandberg, R. 2014. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Research, 24: 2033–2040.

Puncher, G. N., Rowe, S., Rose, G. A., Leblanc, N. M., Parent, G. J., Wang, Y., and Pavey, S. A. 2019. Chromosomal inversions in the Atlantic cod genome: implications for management of Canada's northern cod stock. Fisheries research, 216: 29–40.

Richards, S. W. 1982. Aspects of the biology of *Ammodytes americanus* from the St. Lawrence river to Chesapeake Bay, 1972–75, including a comparison of the long island sound post larvae with *Ammodytes dubius*. Journal of Northwest Atlantic Fishery Science, 3: 93–104.

Richardson, D. E., Palmer, M. C., Smith, B. E., and Cooper, A. 2014. The influence of forage fish abundance on the aggregation of Gulf of Maine Atlantic cod (*Gadus morhua*) and their catchability in the fishery. Canadian Journal of Fisheries and Aquatic Sciences, 71: 1349–1362.

Robards, M. D., Willson, M. F., Armstrong, R. H., and Piatt, J. F. 2000. Sand lance: a review of biology and predator relations and annotated bibliography. Federal Government Series Report, PNW-RP-521. 340 pp.

Silva, T. L., Wiley, D. N., Thompson, M. A., Hong, P., Kaufman, L., Suca, J. J., Llopiz, J. K. *et al.*. 2020. High collocation of sand lance and protected top predators: implications for conservation and management. Conservation Science and Practice, 3: e274.

Smigielski, A. S., Halavik, T. A., Buckley, L. J., Drew, S. M., and Laurence, G. C. 1984. Spawning, embryo development and growth of the American sand lance *Ammodytes americanus* in the laboratory. Marine Ecology Progress Series, 14: 287–292.

Stanley, R. R., DiBacco, C., Lowen, B., Beiko, R. G., Jeffery, N. W., Van Wyngaarden, M., Bentzen, P. *et al.*. 2018. A climate-associated multispecies cryptic cline in the northwest Atlantic. Science Advances, 4: eaaq0929.

Staudinger, M. D., Goyert, H., Suca, J. J., Coleman, K., Welch, L., Llopiz, J. K., Wiley, D. *et al.*. 2020. The role of sand lances (*Ammodytes* sp.) in the northwest Atlantic ecosystem: a synthesis of current knowledge with implications for conservation and management. Fish and Fisheries, 21: 522–556.

Suca, J. J., Baumann, H., Pham, K., Silva, T. L., Wiley, D. N., Feng, Z., and Llopiz, J. K. 2022. Larval transport pathways from three prominent sand lance habitats in the Gulf of Maine. Fisheries Oceanography, 31: 1–20.

Suca, J. J., Wiley, D. N., Silva, T. L., Robuck, A. R., Richardson, D. E., Glancy, S. G., Clancey, E. *et al.*. 2021. Sensitivity of sand lance to shifting prey and hydrography indicates forthcoming change to the northeast US shelf forage fish complex. ICES Journal of Marine Science, 78: 1023–1037.

Therkildsen, N. O., and Palumbi, S. R. 2017. Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. Molecular Ecology Resources, 17: 194–208.

Tremblay, M. J., Loder, J. W., Werner, F. E., Naimie, C. E., Page, F. H., and Sinclair, M. M. 1994. Drift of sea scallop larvae *Placopecten magellanicus* on Georges Bank: a model study of the roles of mean advection, larval behavior and larval origin. Deep Sea Research Part II: Topical Studies in Oceanography, 41: 7–49.

van Deurs, M., van Hal, R., Tomczak, M. T., Jónasdóttir, S. H., and Dolmer, P. 2009. Recruitment of lesser sandeel *Ammodytes marinus* in relation to density dependence and zooplankton composition. Marine Ecology Progress Series, 381: 249–258.

Van Wyngaarden, M., Snelgrove, P. V., DiBacco, C., Hamilton, L. C., Rodríguez-Ezpeleta, N., Jeffery, N. W., Stanley, R. R. *et al.*. 2017. Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RAD seq-derived SNPs. Evolutionary Applications, 10: 102–117.

Wilder, A. P., Palumbi, S. R., Conover, D. O., and Therkildsen, N. O. 2020. Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. Evolution Letters, 4: 430–443.

Winters, G. H., and Dalley, E. L. 1988. Meristic composition of sand lance (*Ammodytes* spp.) in Newfoundland waters with a review of species designations in the northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences, 45: 516–529.

Wright, P., Christensen, A., Régnier, T., Rindorf, A., and van Deurs, M. 2019. Integrating the scale of population processes into fisheries management, as illustrated in the sandeel, *Ammodytes marinus*. ICES Journal of Marine Science, 76: 1453–1463.

Handling Editor: Kristi Miller-Saunders