Mitogenomic evidence of population differentiation of thorny skate, *Amblyraja radiata*, in the North Atlantic

John S. S. Denton[1A], Jeff Kneebone[2AB], Lei Yang[1A], Arve Lynghammar[3], W. David McElroy[4], Shannon Corrigan[1], Klara Jakobsdóttir[5], Carolyn Miri[6], Mark Simpson[6], and Gavin J. P. Naylor[1]

[1]Florida Museum of Natural History, Dickinson Hall, Gainesville, FL, USA

[2]Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA, USA

[3]UiT The Arctic University of Norway, Faculty of Biosciences, Fisheries and Economics, The Norwegian College of Fishery Science, NO-9037 Breivika, Norway

[4]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA, USA

[5]Marine and Freshwater Research Institute, Demersal Division, Hafnarfjörður, Iceland

[6]Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, St John's, Newfoundland, Canada

[A]Authors made equal contributions and should be considered as joint first authors

[B]Corresponding author: Anderson Cabot Center for Ocean Life at the New England Aquarium, 1 Central Wharf, Boston, MA USA 02110, jkneebone@neaq.org

Funding: Lenfest Ocean Program

Abstract

Management of thorny skate (*Amblyraja radiata*) in the Northwest Atlantic has posed a conservation dilemma for several decades due to the species' lack of response to strong conservation efforts in the U.S. Gulf of Maine and Canadian Scotian Shelf, confusion over the relationship between two reproductive size morphs of differing life histories that are sympatric in the Northwest Atlantic, and conflicting data on regional population connectivity throughout the species' broader range. To better assess potential *A. radiata* regional population differentiation and genetic links to life-history variation, we analyzed complete mitochondrial genome sequences from 527 specimens collected across the species' complete North Atlantic geographic range, with particular emphasis on the Northwest Atlantic region. A high level of genetic diversity was evident across the North Atlantic, but significant genetic differentiation was identified between specimens inhabiting the Northwest (Gulf of Maine and Newfoundland) and Northeast (Greenland, Iceland, North Sea, Arctic Circle) Atlantic. In the Northwest Atlantic, significant differentiation between the Gulf of Maine and Newfoundland regions was revealed, however, the overall level of differentiation was very low. No genetic difference was identified between the large and small reproductive morphs. The results of this study advance our understanding of *A. radiata* population structure in the North Atlantic, but do not resolve all of the questions confounding our understanding of the species' biology and evolutionary history.

Keywords: genetics, mitogenome, population structure, starry skate, starry ray

Introduction

The thorny skate (*Amblyraja radiata*), also known as the starry skate or starry ray (Donovan 1808), is a wide-ranging North Atlantic species, with a geographic range from South Carolina (USA) north to Baffin Bay (Arctic Ocean) in the Northwest Atlantic, and from the coasts of Greenland, Iceland, Ireland, England north to Norway, Kattegat in the North Sea, the Svalbard Archipelago in the Barents Sea, and as far east as Novaya Zemlya (Bigelow and Schroeder, 1953; Mecklenburg *et al.,* 2018). Recent Bayesian population analyses across this range identified a strong latitudinal pattern in *A. radiata* population status over the past five decades, with steep declines (>80%) evident at the most southerly latitudes and stable and/or increasing abundance at more northerly latitudes (Kulka *et al.,* 2020). This gradient of population trends is of significant interest to fishery managers and population biologists, particularly in relation to the evaluation of existing and future threats to *A. radiata* from fishing pressure and climate change (e.g., Grieve *et al.,* 2020).

Identifying causes of the disparate *A. radiata* population trends across the North Atlantic is complicated by an incomplete understanding of the species' population dynamics and life history throughout this vast region. For example, *A. radiata* exhibit highly variable sizes at maturity throughout the North Atlantic, and at least two size morphs (that mature at distinctly different sizes and ages) occur off Canada and the United States in the Northwest Atlantic (Templeman, 1987a; Templeman, 1987b; Sosebee, 2005; Sulikowski *et al.,* 2005; McPhie and Campana, 2009). Interestingly, no genetic difference between these size morphs has been detected (Ostrow *et al.,* 2008; Lynghammar *et al.,* 2016). Genetic analyses of *A. radiata* sampled throughout the North Atlantic have also yielded inconsistent results on population structuring. Chevolot *et al.* (2007) sequenced a 290 bp fragment of the cytochrome *b* (Cyt *b*) gene and found

68 little evidence for genetic differentiation across Newfoundland, Iceland, Norway, the North Sea,
69 and the Kattegat, with generally high haplotype diversity outside the North Sea. In contrast,
70 analyses of a 651 bp fragment of the cytochrome oxidase subunit 1 (COI) gene (Lynghammar *et*
71 *al.,* 2014) and 10 microsatellite loci (Lynghammar *et al.,* 2016) revealed evidence of divergence
72 across the North Atlantic, with slight support for three major geographic groupings (Northwest
73 Atlantic, Greenland, and Northeast Atlantic). Tagging studies also demonstrate that *A. radiata*
74 exhibit a high degree of residency to specific regions in the North Atlantic, and generally
75 undergo horizontal displacements of <100 km over as long as 20 years at liberty (Templeman,
76 1984; Walker *et al.,* 1997; Kneebone *et al.,* 2020). This lack of movement/dispersal suggests the
77 potential for spatial structuring, differentiated stocks, or cryptic speciation. However, genetic
78 data collected thus far do not corroborate the degree of structuring suggested by life history and
79 phenology.
80 Given the inconsistencies among the available data on *A. radiata* life history, population
81 genetics, and population status, management of the species in the North Atlantic remains a
82 conundrum. It may be that genomic and spatial sampling has so far been insufficient to capture
83 subtle signals of biologically relevant variation that explain observed patterns (Costa *et al.,* 2015)
84 since most previous genetic analyses have not included samples from the species' full
85 geographic range, particularly U.S. waters of the Gulf of Maine and Georges Bank (Chevolot *et*
86 *al.,* 2007; Lynghammar *et al.,* 2014, 2016). The most comprehensively sampled locus in *A.*
87 *radiata* analyses to date—the mitochondrion—also exhibits a heterogeneous process history,
88 where different pressures have been exerted in different regions, resulting in a variety of
89 signatures (Rubinoff *et al.,* 2006). Analyzing complete mitochondrial genomes (mitogenomes)
90 has been shown to improve parameter estimates in phylogenetic studies relative to single

fragments (Cummings *et al.,* 1995; Hancock-Hanser *et al.,* 2013, and sources therein), and pipeline optimizations and cost reductions for next-generation sequencing make sampling both broadly and deeply increasingly feasible for non-model organisms such as *A. radiata*.

The goal of this study was to sequence entire mitogenomes from *A. radiata* sampled over a broad geographic range to better assess potential population differentiation and genetic links to life-history variation across the North Atlantic, with a particular emphasis on the Northwest Atlantic region. This region is of heightened management interest due to the severe *A. radiata* population declines evident from the Gulf of Saint Lawrence south to the Gulf of Maine (Kulka *et al*., 2020) and the unique co-existence of the large and small size morphs. Population decline in the U.S. Gulf of Maine and Canadian Scotian Shelf has already led to fishery closures (NEFMC, 2003; COSEWIC, 2012) and multiple petitions to list *A. radiata* as Endangered under the U.S. Endangered Species Act, the most recent of which argued that individuals in the Northwest Atlantic exist as a distinct population segment. Uncertainty over the biological connectivity between the large and small size morphs also continues to confound understanding of the mechanisms contributing to the lack of recovery in U.S. waters despite 20 years of fishery closures. Accordingly, we sought to expand upon the work of Lynghammar *et al.* (2016) by conducting more extensive biological sampling throughout the Northwest Atlantic region to (1) further examine the extent to which *A. radiata* in this region may be differentiated from other regions of the North Atlantic, and (2) re-assess genetic differentiation between the large and small size morphs that occur in the U.S. Gulf of Maine.

Materials and Methods

Ethical statement

No live animals were intentionally sampled by this project. Instead, archived tissue samples collected as part of fisheries surveys conducted throughout the North Atlantic by various countries were obtained for analysis. As such, no ethical approval was required.

Specimen collection and classification

Tissue samples were obtained from 527 individual *A. radiata* captured throughout the species' North Atlantic range (Figure 1). All tissue samples were collected during fisheries surveys, preserved in 95% ethanol, and sent to the Florida Museum of Natural History (Gainesville, Florida, USA) for processing. See Table S1 for the full sample information. To facilitate the analysis of population structure throughout the North Atlantic, samples were assigned a geographic region *a priori sensu* Lynghammar *et al.* (2016). Nine geographic regions were identified: Gulf of Maine (GoM), Newfoundland (NL), South Greenland (SG), East Greenland (EG), West Iceland (WI), East Iceland (EI), North Sea (Nsea), Northern Norway (Nnor), and Arctic Circle (ArcC) (Figure 1). Note that abbreviations are used for these regions when referenced to the analysis or results, but region names are spelled out when discussed in a broader context.

Library preparation and target enrichment

Total genomic DNA was extracted using the E.Z.N.A. Tissue DNA kit (Omega Bio-Tek, Norcross, Georgia, USA) per the manufacturer instructions. Aliquots of extracted genomic DNA (0.5 – 3.0 µg) from each sample were sheared to a mean fragment length of 500 bp using a Covaris M220 ultrasonicator (Covaris, Inc., Massachusetts, USA). Dual-indexed genomic libraries suitable for sequencing on an Illumina platform were prepared from the fragmented DNA following the methods of Meyer and Kircher (2010) and Fisher *et al.* (2011), as detailed in

137 Li *et al.* (2013). Library size selection steps were one-sided, removing fragments smaller than
138 200 bp, and conducted via a solid phase reversible immobilization (SPRI) protocol modified
139 from Rohland and Reich (2012), using Carboxyl-modified Sera-Mag Magnetic Speed-beads
140 (Fisher Scientific, Waltham, Massachusetts, USA) in a 'MagNA' (PEG/NaCl) buffer. Final
141 amplified and indexed libraries were eluted into 20 μL of nuclease-free water. The quality and
142 length distribution of the library was visualized using gel electrophoresis.
143 Separate enrichment experiments were performed for each sample library using a custom
144 biotinylated RNA probe set (MYbaits MYcroarray, Ann Arbor, Michigan, USA) that was
145 designed to target the complete mitochondrial genome of skates by including bait sequences
146 derived from 83 species that span the known diversity of batoids (White *et al.,* 2018).
147 Enrichment followed the relaxed hybridization method described by Li *et al.* (2013).
148 Mitogenome enriched sample libraries were pooled in equimolar ratios. Quality and quantity of
149 the pooled library was assessed via qPCR and using the D5000 ScreenTape System (Agilent,
150 Santa Clara, California, USA) and then diluted to 8 pM for paired-end 300 bp sequencing on an
151 Illumina MiSeq sequencer at the University of Florida Interdisciplinary Center for
152 Biotechnology Research.
153
154 Mitogenome assembly and annotation
155 Adaptors, indexes, and low-quality sequences were batch trimmed from demultiplexed
156 sequences via paired-read calls to Trimgalore v0.6.4 + cutadapt (Krueger, 2019) using custom
157 bash scripts. Trimmed reads for each sample were imported as sample-specific paired read lists
158 into Geneious v.11.1.5 (https://www.geneious.com). Duplicated reads were removed from each
159 paired read list, and unique reads were assembled as circular genomes to a previously sequenced

A. radiata mitogenome (GN2602; GenBank accession # OP651989), which was annotated using the MitoAnnotator pipeline for fish mitogenomes (Iwasaki *et al.,* 2013). The manually curated GN2602 annotations were used as a library for automatic re-annotation of newly sequenced *A. radiata* mitogenomes. The removal of duplicated reads, mitogenome assembly, and annotation were accomplished via a custom-written Geneious batch workflow. For consensus sequence generation in Geneious we set the threshold as "Highest quality: use chromatogram quality to call the best base". Sequences of each protein-coding gene were concatenated into the final alignment. The ND6 gene was reverse-complemented before concatenation. All stop codons (both partial and complete) and alignment-ambiguous mitogenomic regions (12S, 16S, control region, and tRNA elements) were removed. All mitogenome sequences generated in this study have been submitted to GenBank under the accession numbers OR812525-OR813050 (Table S1).

Population structure and differentiation

To visualize haplotype clustering and distribution, a haplotype network was inferred from protein-coding mitogenome sequences using the TCS network method (Templeton *et al.,* 1992) in popART (Leigh and Bryant, 2015). A rooted phylogenetic tree, depicting relative branch lengths, was also inferred using maximum likelihood based on 443 unique protein-coding mitogenome sequences/haplotypes identified across all samples. Two species, *Amblyraja doellojuradoi* (GN2381; mitogenome GenBank accession # OP651988) and *Rajella fyllae* (GN2135; mitogenome GenBank accession # OP651990), were added as outgroups. The genus *Rajella* was included in the analyses because it was identified as the sister taxa of *Amblyraja* (Naylor *et al.,* 2012), which is important for later timetree analyses. PartitionFinder 2 (Lanfear *et*

183 *al.,* 2017) was used to find the best partitioning scheme (15 partitions) and nucleotide
184 substitution model (GTR + I + G) for the mitogenome alignment. Partitioned Maximum
185 Likelihood (ML) analysis was conducted using RAxML v.8.0.26 (Stamatakis, 2006, 2014). Two
186 hundred independent runs were performed based on 200 random starting trees using the default
187 algorithm of the program. The tree with the highest likelihood score (*Ln*L) was chosen as the
188 final tree. Bootstrap analyses (1000 replicates) were also conducted using RAxML (Felsenstein
189 1985; Stamatakis *et al.,* 2008) with the same partitioning strategy and nucleotide substitution
190 model as above. PAUP 4.0.b10 (Swofford, 2002) was then employed to obtain the 50% majority
191 rule consensus tree and bootstrap values (BP).
192 To assess genetic structure and differentiation among *A. radiata* sampled throughout the
193 North Atlantic, an analysis of molecular variance (AMOVA) was conducted in Arlequin 3.5
194 (Excoffier and Lischer, 2010). Given our objective to better understand genetic/population
195 structuring between *A. radiata* of the Northwest Atlantic and the rest of the species' range, the
196 GoM and NL samples were consolidated into a single group (Northwest) and compared to a
197 'Northeast' group that comprised the seven remaining regions (i.e., SG, EG, WI, EI, ArcC, Nsea,
198 and Nnor). The significance test was assessed using 1000 permutations. To further explore the
199 population differentiation between sampled regions in the North Atlantic for which at least 10
200 samples were analyzed, the pairwise F_{ST} value was calculated for each pair. The pairwise
201 distance (p-distance) was used and the significance test was assessed using 1000 permutations.
202 The significance level was set as $p < 0.01$ for all analyses.
203
204 <u>Timetree analyses and ancestral range reconstruction</u>

205 Timetree analyses were performed to better explore the processes that might have
206 contributed to the current distribution of *A. radiata* in the North Atlantic. As there are no known
207 *A. radiata* fossils, the time tree was calibrated using the divergence time estimates (7.56-18.97
208 million years ago; mya) for the node leads to the sister genera *Amblyraja* and *Rajella* from
209 Aschliman *et al.* (2012), which itself used many carefully selected fossils to calibrate a timetree
210 across batoids. To time-calibrate the phylogenetic tree using this *Amblyraja-Rajella* node and
211 produce a distribution of calibrated trees, we sampled 500 random numbers/ages following a
212 lognormal distribution with the following parameters: offset = 7.56, mean = 1.612, standard
213 deviation = 0.5, and 95% distribution = 18.97. We then conducted separate timetree analyses in
214 TreePL (Smith and O'Meara, 2012) using 7.56 mya as the minimum age and each of the 500
215 numbers generated above as the maximum age for the root node. For each individual TreePL
216 analysis, we first performed cross-validation analyses using the "cv" and "randomcv" commands.
217 We then tested the performance of the available optimization routines using the "prime"
218 command. Lastly, we used the "thorough" command to ensure the preferred optimization routine
219 converged. The 500 resulting timetrees were combined into a single 500-tree Newick file and
220 imported into TreeAnnotator 2.7.1 (Bouckaert *et al.,* 2019). A maximum clade credibility (MCC)
221 tree was then calculated (no burn-in; posterior probability limit 0.0; median node heights).
222 The topologies of the phylogenetic tree and the haplotype network suggested *A. radiata*
223 might have experienced multiple East-West colonization events in the past. To test this
224 hypothesis, we reconstructed the ancestral biogeographic distribution of *A. radiata* using
225 BioGeoBEARS (Matzke, 2013). The MCC tree generated by our divergence time analyses was
226 used as the input phylogeny after the two outgroups were removed. We defined nine
227 biogeographic areas (A - I) that correspond to the nine geographical populations (GoM, NL, SG,

228 EG, WI, EI, Nsea, Nnor, and ArcC), respectively. Because most of the nine geographical areas
229 are far from one another and *A. radiata* have low potential for long range dispersal
230 (Lynghammar *et al.,* 2016), the maximum number of ancestral areas allowed at each node was
231 set to four. According to Jørgensen *et al.* (2005), *A. radiata* are obligate bottom-dwellers down
232 to a depth of 1,442 m. They are more likely to disperse along the continental shelves than across
233 the deep ocean separating the East Greenland and the Arctic Circle areas and the deep ocean
234 separating the Northwest Atlantic region and the Iceland area (Figure 1; Lynghammar *et al.,*
235 2016). Therefore, distances among the nine areas were measured along the continental shelves
236 using the shortest distance between the assumed centers of distribution. All measurements were
237 performed using tools implemented in Google Earth Pro version 9.3.6. Following the suggestion
238 from the author of BioGeoBEARS, the distance matrix was re-scaled by dividing all the
239 distances by the shortest distance, so that the matrix only contains values larger or equal to one.
240 All values in the matrix were also rounded to the first decimal digit. The re-scaling and rounding
241 can greatly reduce the impact of measurement uncertainty. Connectivity of biogeographic areas
242 was modeled with the following two dispersal probability categories: 1.0 for well-connected
243 areas (e.g., GoM and NL), and 0.0001 for well-separated areas (e.g., East Greenland and Arctic
244 Circle). Six models (DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, and
245 BAYAREALIKE+J) implemented in BioGeoBEARS were tested. Statistical significance was
246 accepted at $p < 0.001$.
247
248 Demographic analyses
249 The summary statistic Fu's *Fs* (Fu, 1997) was estimated for each of the nine geographical
250 regions, the Northeast group, and the Northwest group using Arlequin 3.5 (Excoffier and

Lischer, 2010) to test for departure from a constant population size. The *p*-distance was used and statistical significance was assessed using 1000 permutations. We did not estimate Fu's *Fs* for the entire sample because it is not panmictic. Statistical significance was accepted at $p < 0.001$.

Population structure of size morphs

To investigate potential genetic differentiation between the large (late maturing) and small (early maturing) size morphs of *A. radiata* found in the GoM, all specimens from this region with size, sex, and reproductive status data (n=148) were categorized as "large" or "small" according to conditional reasoning. Briefly, all mature specimens that measured <75 cm total length (i.e., ~10 cm smaller than the size at 50% maturity reported for both sexes in Sulikowski *et al.,* 2006) were considered members of the small morph. Alternatively, all mature individuals measuring >85 cm total length (the reported size at 50% maturity; Sulikowski *et al.,* 2006) as well as all immature individuals measuring 55 – 75 cm total length (i.e., the size range over which members of the small morph are reproductively active; Sosebee, 2005) were assigned to the large morph.

To assess the presence of interaction between sex and geographic region and sex and size morph, two non-nested two-way AMOVA were conducted, using the method of Iwaizumi *et al.* (2013). The first two-way AMOVA compared sex × region, for all samples with both sex and region information available (N = 482). The second two-way AMOVA examined sex × size morph in the Gulf of Maine. In this second AMOVA, statistical significance was assessed using 1000 permutations and the significance level of each AMOVA was set as $p < 0.01$.

Results

Sequence characteristics and diversity measurements

The complete *A. radiata* mitogenome (16,795 bp), sequenced at 400x coverage, exhibited the typical vertebrate arrangement, with 13 protein-coding genes (11,391 bp), 22 tRNA genes, 2 rRNA genes, and 1 control region. The protein-coding alignment, on which analyses were based, included 1,339 variable sites and 782 parsimony-informative sites. There were 443 unique haplotypes among the 527 sampled individuals. Global nucleotide and haplotype diversities were 0.0062 and 0.9992, respectively, while regional nucleotide diversities ranged between 0.0037 and 0.0061 (Table 1). Regional haplotype diversities ranged between 0.9455 and 1.0000.

Population structure and differentiation

The haplotype network (Figures 2 and S1) was characterized by high diversity with little haplotype sharing among individuals or regions. Clade 1 was composed almost exclusively (99.1%) of GoM/NL haplotypes, but included one EI haplotype and one WI haplotype. Haplotypes in clade 1 were usually separated by only one or a few steps. Clade 2 was composed of a complicated network of haplotypes representing all nine regions. Many haplotypes are separated by > 5 steps. Clade 3 contained only six haplotypes (1 NL, 1 SG, 1 ArcC, and 3 EI). Three major clades (1-3) were resolved from the reconstructed phylogenetic tree (Figure S2), all of which were robustly supported (BP = 100%).

One-way hierarchical AMOVA showed most genetic variation occurred within geographic regions (Table 2; Vc = 29.527; df = 518; percentage variance = 68.93%). Over a quarter of genetic variation occurred between the Northwest Atlantic (GoM & NL) and other regions (Va = 12.290; df = 1; percentage variance = 28.69%). Both F_{ST} (0.31070; $p < 0.0001$) and F_{SC} (0.03337; $p < 0.0001$) were significant. F_{CT} (0.28690; $p = 0.02444$) was not significant.

297 Pairwise F_{ST} values for geographic regions (Table 3) recovered a pattern similar to that observed
298 in the TCS network. The GoM and NL samples were significantly different from all other
299 regions, including each other (F_{ST} = 0.0556; p = 0.00098). No other geographic regions exhibited
300 significant pairwise differences in F_{ST}.
301
302 Timetree analyses and ancestral range reconstruction
303 Timetree analyses indicated that divergences among *A. radiata* sampled in this study
304 occurred in the past one million years (Figure 3 and S3). Clade 1 began to diverge at 0.36 million
305 years ago (mya), clade 2 at 0.67 mya, and clade 3 at 0.11 mya. BioGeoBEARS models with the
306 parameter "J" consistently returned higher likelihood values and included some instances of
307 long-distance dispersals ("jumps") that are not feasible for *A. radiata*, such as direct dispersals
308 between the Arctic and the Gulf of Maine. As a result, results obtained from the three models
309 with "J" were disregarded and are not presented. BioGeoBEARS analysis based on the
310 DIVALIKE model returned a higher likelihood value (*Ln*L=−1098.98) than the DEC
311 (*Ln*L=−1123.64), and BAYAREALIKE (*Ln*L=−1307.52) models, and thus were interpreted
312 further. The ancestral range reconstruction results suggested that there might have been as many
313 as 28 east-west colonization events and only four west-east colonization events within the last
314 0.67 million years (Figure 3 and S4). The four West-East colonization events also seem to have
315 happened recently (< 0.3 mya).
316
317 Demographic analyses

318 Negative and significant Fu's *Fs* values were found for the Northwest group, Northeast
319 group, and the Gulf of Maine, indicating that historical population expansion events might have
320 happened in these samples (Table 1).
321
322 Population structure of size morphs
323 Two-way AMOVA of sex × region recovered a significant effect of region (R^2 =
324 0.28013; p = 0.001), but not sex (R^2 = 0.00259; p = 0.156) (Table 4). There was no significant
325 interaction between sex and region (R^2 = 0.00993; p = 0.671). Two-way AMOVA of sex × size
326 morph in the GoM recovered no significant main effects (sex: R^2 = 0.00339; p = 0.589; morph:
327 R^2 = 0.00182; p = 0.823) or interaction (R^2 = 0.00147; p = 0.899) (Table 5). After mapping size
328 morph information on the haplotype network (Figure S5), we noticed that there were several
329 cases of haplotype sharing between the large morph and the small morph. We also noticed that
330 there were no relatively large clusters formed by any single size morph.
331
332 **Discussion**
333 This study builds upon previous work on *A. radiata* population structure and historical
334 demography in the North Atlantic (e.g., Chevlot *et al.,* 2007; Coulson *et al.,* 2011; Lynghammar
335 *et al.,* 2014, 2016) using whole mitochondrial genomes and a more comprehensive sampling
336 scheme that includes a large number of samples from U.S. waters, a region of intense
337 conservation and management interest (e.g., NMFS, 2017) that was not well sampled in previous
338 studies. Consistent with Lynghammar *et al.* (2016), our results provide evidence that *A. radiata*
339 in the Northwest Atlantic are distinct from those occurring off Greenland, Iceland, and the
340 broader Northeast Atlantic. We also found statistical support for genetic differentiation between

individuals inhabiting the GoM and NL regions within the Northwest Atlantic, however, the level of differentiation was very low (F_{ST} = 0.0556). No evidence of genetic differentiation was observed between skates of the large and small reproductive morphs that were sampled in the U.S. waters of the GoM. Collectively, the results of this study advance our understanding of *A. radiata* population structure in the North Atlantic, but do not resolve all of the questions confounding our understanding of the species' biology and evolutionary history throughout its vast range.

Sequencing of full *A. radiata* mitogenomes revealed a high level of genetic diversity across the North Atlantic, as evidenced by the occurrence of 443 unique protein-coding haplotypes among the 527 sampled individuals. High genetic diversity has been previously demonstrated by multiple studies of *A. radiata* population structure in the North Atlantic (e.g., Chevolot *et al.,* 2007; Coulson *et al.,* 2011; Lynghammar *et al.,* 2014, 2016), and has been hypothesized to occur due to widespread physical mixing of individuals throughout the species' range. However, *A. radiata* demographics and direct measurement of movement patterns through tagging neither support high rates of movement nor support broad-scale migrations between any of the nine regions sampled (Templeman, 1984; Walker *et al.,* 1997; Kneebone *et al.,* 2020). *Amblyraja radiata* is also an obligate bottom-dweller that exists almost exclusively in continental shelf waters from 25 to 440 m (Kulka *et al.,* 2020) and to depths up to 1,442 m (Jørgensen *et al.,* 2005), which restricts available migration pathways to the continental shelf waters and precludes direct movement between distant regions separated by deep ocean habitats. Nonetheless, phylogenetic and biogeographical analyses performed by this and previous studies suggest there have been numerous east-to-west *A. radiata* population expansions in the North Atlantic from 0.36 to 0.67 mya (this study) and 0.6-1.1 mya (Chevolot *et al.,* 2007). It is possible that physical

364 mixing occurred during periods of glacial expansion and retreat. However, given the prevalence
365 of sedentary behavior and low dispersion in contemporary populations, it is likely that high
366 levels of genetic diversity and connections among distant regions are remnants of ancestral
367 colonization events rather than active contemporary mixing (Lynghammar *et al.,* 2016).

368 Pairwise F_{ST} analyses revealed variable levels of genetic diversity among regions of the
369 North Atlantic for which at least 10 samples were available, with greater divergence generally
370 evident between populations separated by large geographic distances. Diversity was lowest
371 among the regions encompassing the Northeast grouping, with no significant differentiation
372 evident between any region. However, significant divergence was evident between each of the
373 Northeast regions and the GoM, and between NL and all Northeast regions except SG. Of note,
374 only 11 samples were analyzed from SG, thus, these results should be treated with caution. No
375 previous studies on *A. radiata* North Atlantic population genetics have included samples from
376 the GoM, but similar results were reported by Lynghammar *et al.* (2016) who found significant
377 differentiation between Newfoundland and Labrador, Canada, and the ArcC, Nnor, Nsea, and
378 SG, but not EG, based on an assessment of 10 microsatellite loci. In this study, the NL and SG
379 regions were not found to be significantly differentiated, which suggests that Greenland was, and
380 may still be, a conduit for gene flow between *A. radiata* inhabiting the eastern and western
381 extents of the North Atlantic.

382 Significant divergence was evident between the GoM and NL regions based on pairwise
383 F_{ST} analyses, despite the low level of differentiation ($F_{ST} = 0.0556$). Haplotype networks of GoM
384 and NL samples in both clades 1 and 2 were also often only separated by only one or a few steps
385 and there was evidence of reciprocal gene flows between them (GoM to NL and NL to GoM),
386 which further suggests connectivity between these regions despite the apparent lack of

movement between them on a decadal time scale (Templeman, 1984; Kneebone *et al.,* 2020). Coulson *et al.* (2011) previously reported ~1% COI barcode divergence between *A. radiata* sampled in the Northwest Atlantic and all individuals clustered together, including a single specimen from the U.S. GoM that exhibited 3-4% COI divergence. Unpublished analyses of five microsatellite loci also showed no genetic structuring among 95 specimens of *A. radiata* sampled at four locations in the Gulf of Maine (two locations; n = 60) and Canada (two locations; n = 35) and revealed 98% of the genetic variation was explained by within population variation, while less than 2% could be attributed to variation among different populations (Tsang *et al.,* 2008). Curiously, however, our study revealed no shared haplotypes between the GoM and NL regions, suggesting a lack of widespread interbreeding. No hydrological barriers (e.g., deep channels >1,400 m) exist that would impede gene flow between the GoM and NL regions. Thus, it's possible that gradual mixing of *A. radiata* that are distributed across the Scotian Shelf (NAFO division 4) serves as the link between the GoM and NL regions. Nonetheless, given the large haplotype diversities inferred for the *A. radiata* relative to other skate species (e.g. Coulson *et al.,* 2011), it is possible that additional genetic substructure could emerge with larger and regionally finer-scale sampling in the Northwest Atlantic, including the Scotian Shelf and Gulf of St. Lawrence, areas which span major hydrological, physical, and oceanographic features (Han *et al.,* 1999), and which are currently unsampled.

The lack of genetic differentiation between the large and small reproductive morphs that were sampled in the GoM adds to the growing conundrum over the evolutionary history of *A. radiata* in the Northwest Atlantic and the source of such a large discrepancy in life history between sympatric individuals of the same species. Based on their marked differences in size-at-maturity and maximum size (Templeman, 1987a; Sulikowski *et al.,* 2005; Sosebee, 2005) mating

between individuals of the large and small morphs was considered unlikely. Curiously, no studies have found any genetic differentiation between individuals from each size morph based on the COI gene (Lynghammar *et al.,* 2014), full mitochondrial genomes (this study), or multiple microsatellite loci (Ostrow *et al.,* 2008; Tsang *et al.,* 2008; Lynghammar *et al.,* 2016). *Amblyraja radiata* representing the large and small morphs are also sympatric, exhibit similar horizontal movements, and occur in identical habitats (e.g., substrate, depth, temperature) in the GoM (Kneebone *et al.,* 2020), thus, environmental effects on gene expression are unlikely the cause of such drastic differences in life history. Clearly, questions of morphological and ecological correlates have important implications for management, but even with complete mitogenomic data the present results do not indicate clear relationships between the life-history traits and genetic variation.

We note, however, that even though the current sampling is geographically comprehensive, and the amount of data collected (whole mitochondrial genomes) more extensive than any prior studies on thorny skates, our inferences are nevertheless based on a non-recombining, single locus "super-gene." Although this super-gene provides better resolution of mitochondrial evolution than do individual mitochondrial sub-fragments, its utility ultimately relies on how well the inference obtained captures demographic and evolutionary history. Whole mitochondrial genome analysis is useful for reconstructing evolutionary histories within populations because mitogenomes are faster-evolving, and have smaller effective population sizes, than nuclear genomes. These features enable mitogenomes to have greater resolving power than nuclear genes for recent divergences, such as at the subpopulation level. However, due to the matrilineal inheritance pattern of mitochondrial DNA, historical signatures inferred from the data reflect this inheritance pattern and the evolution of the mitochondrion itself, which may not

433 align with the evolution of the organism, as for example in cases of introgression between
434 lineages. By contrast, approaches like short fragment analysis of nuclear genome data is bi-
435 parentally inherited and so do not suffer from a matrilineally biased inheritance pattern.
436 However, the lengths of short reads make assignment of homology among them difficult, even
437 when mapping to a reference genome, when genomic elements have been duplicated. Short read
438 nuclear data are also generally less fast evolving than are mitochondrial data. This fact, together
439 with the larger effective population size of nuclear relative to mitochondrial data, means more
440 short read markers are required to obtain estimates of comparable resolution, which can be
441 challenging when resources are limited. Despite these limitations, given the complex history of
442 the thorny skate, for which processes such as lineage sorting and introgression must be
443 accounted, nuclear data should be collected to enable generation of a consensus over a broad
444 range of gene histories and to circumvent the potential biases associated with reliance on single
445 gene inferences such as those based on the mitogenome.
446 The results of this study provide strong evidence of *A. radiata* genetic structuring across
447 the North Atlantic and confirm the assertion of Lynghammar *et al.* (2016) that regional
448 population management is required for species management. In the Northwest Atlantic, the low,
449 albeit statistically significant, level of differentiation between *A. radiata* inhabiting the GoM and
450 NL regions indicates some degree of gene flow occurs over this broad geographic area, and that
451 coordinated management between the U.S. and Canada remains necessary. This is particularly
452 warranted at more southerly latitudes of the Gulf of Maine and Scotian Shelf where perpetual
453 population decline has already resulted in fishery restrictions over the past two decades
454 (NEFMC, 2003; COSEWIC, 2012) and abundance remains low (Kulka *et al.,* 2020). Given the
455 documented negative effects of climate change and warming ocean conditions on *A. radiata*

456 biomass in the North Sea (Sguotti *et al.,* 2016), Gulf of Maine (Nye *et al.,* 2009; Hare *et al.,*
457 2016; Grieve *et al.,* 2020), and Canadian Scotian Shelf (Stortini *et al.,* 2015; Jubinville *et al.,*
458 2021), additional proactive management may be warranted in the Newfoundland region (e.g.,
459 NAFO division 3) where an increase in abundance (8.5%) has occurred in the past three
460 generation lengths despite the existence of a directed fishery for the species on the Grand Banks
461 (Kulka *et al.,* 2020). Proactive management measures may be required in this region to ensure
462 the persistence and prosperity of the genetically distinct Northwest Atlantic population of *A.*
463 *radiata*. Additional regional management measures, such as habitat or time-area closures and
464 bycatch avoidance policies (e.g., Jubenville *et al.,* 2021), should also be considered to ensure
465 fishery sustainability and/or species persistence, particularly in regions where strong declines are
466 evident (e.g., Gulf of Maine, Scotian Shelf, North Sea).
467
468 **Acknowledgements**
469 We thank the Norwegian Institute of Marine Research and the NOAA Northeast Fisheries
470 Science Center Bottom Longline Survey staff, in particular C. Alexander, B. Gervelis, G.
471 Gianesin, and D. St. Amand, and Captains P. Lynch (F/V Mary Elizabeth) and E. Hesse (F/V
472 Tenacious II) and their crews, for sample collection. We are grateful to D. Richardson and staff
473 of the Ecosystems Surveys Branch for providing samples from the NEFSC bottom trawl survey.
474 J. Y. Poulsen at the Greenland Institute for Natural Resources, H. Fock at the Thunen-Institut für
475 Seefischerei and the TUNU-Programme at UiT The Arctic University of Norway kindly
476 provided samples from Greenland. N. Aschliman is thanked for providing necessary data for the
477 timetree analyses. We thank the University of Florida Interdisciplinary Center for Biotechnology
478 Research and their staff for assistance with sequencing. Funding was provided by the Lenfest

Ocean Program. Mention of products in this study is for descriptive purposes and does not indicate endorsement by NOAA Fisheries.

Contributions: G.N., J.K., A.L. conceived the idea for the study. A.L., K.J., C.M., J.K., and W.D.M. provided tissue samples. J.D. and L.Y. collected data and performed the analysis. J.D., J.K., and L.Y. wrote the manuscript. J.K., L.Y., A.L., W.D.M., S.C., and G.N. reviewed and commented on manuscript drafts.

Literature Cited

Aschliman, N.C., Nishida, M., Miya, M., Inoue, J.G., Rosana, K.M., & Naylor, G.J.P. (2012) Body plan convergence in the evolution of skates and rays (Chondrichthyes: Batoidea). *Molecular Phylogenetics and Evolution,* **63**, 28–41.

Bigelow, H.B. & Schroeder, W.C. (1953) Fishes of the western North Atlantic. Part two. Sawfishes, guitarfishes, skates and rays. *Memoirs of the Sears Foundation of Marine Research,* **1,** 1–514.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., Fourment M., Gavryushkina A., et al. (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology,* **15,** e1006650.

Chevolot, M., Wolfs, P.H., Pálsson, J., Rijnsdorp, A.D., Stam, W.T. & Olsen, J.L. (2007) Population structure and historical demography of the thorny skate (*Amblyraja radiata*, Rajidae) in the North Atlantic. *Marine Biology,* **151,** 1275–1286.

COSEWIC. (2012) COSEWIC assessment and status report on the Thorny Skate *Amblyraja radiata* in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa, xviii + 75 pp.

Costa, L.S.d., Corneleo, N.S. & Stefenon, V.M. (2015) Conservation of Forest Biodiversity: how sample size affects the estimation of genetic parameters. *Anais da Academia Brasileira de Ciencias,* **87,** 1095–1100.

Coulson, M., Denti, D., Van Guelpen, L., Miri, C., Kenchington, E. & Bentzen, P. (2011) DNA barcoding of Canada's skates. *Molecular Ecology Resources,* **11,** 968–978.

Cummings, M.P., Otto, S.P. & Wakeley, J. (1995) Sampling properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution,* **12,** 814–822.

510 Excoffier, L. & Lischer, H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform
511 population genetics analyses under Linux and Windows. *Molecular Ecology Resources,*
512 **10,** 564–567.
513 Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap.
514 *Evolution,* **39**, 783–791.
515 Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T.M., Young, G., Fennell, T.J.,
516 Allen, A. & Ambrogio, L. (2011) A scalable, fully automated process for construction of
517 sequence-ready human exome targeted capture libraries. *Genome Biology,* **12,** R1.
518 Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth,
519 hitchhiking and background selection. *Genetics,* **147**, 915–925.
520 Grieve, B.D., Hare, J.A. & McElroy, W.D. (2021) Modeling the impacts of climate change on
521 thorny skate (*Amblyraja radiata*) on the Northeast US shelf using trawl and longline
522 surveys. *Fisheries Oceanography*, ***30***(3), 300–314.
523 Han, G., Loder, J.W. & Smith, P.C. (1999) Seasonal-mean hydrography and circulation in the
524 Gulf of St. Lawrence and on the eastern Scotian and southern Newfoundland shelves.
525 *Journal of Physical Oceanography,* **29,** 1279–1301.
526 Hancock-Hanser, B.L., Frey, A., Leslie, M.S., Dutton, P.H., Archer, F.I. & Morin, P.A. (2013)
527 Targeted multiplex next-generation sequencing: Advances in techniques of mitochondrial
528 and nuclear DNA sequencing for population genomics. *Molecular Ecology Resources,*
529 **13,** 254–268.
530 Hare, J.A., Morrison, W.E., Nelson, M.W., Stachura, M.M., Teeters, E.J., Griffis, R.B.,
531 Alexander, M.A., Scott, J.D., Alade, L., Bell, R.J., Chute, A.S. *et al.* (2016) A

vulnerability assessment of fish and invertebrates to climate change on the Northeast US Continental Shelf. *PloS one*, **11**(2), p.e0146756.

Iwaizumi, M.G., Takahashi, M., Isoda, K. & Austerlitz, F. (2013) Consecutive five-year analysis of paternal and maternal gene flow and contributions of gametic heterogeneities to overall genetic composition of dispersed seeds of *Pinus densiflora* (Pinaceae). *American Journal of Botany,* **100,** 1896–1904.

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T.P., Sado, T., Mabuchi, K., Takeshima, H. & Miya, M. (2013) MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Molecular Biology and Evolution,* **30,** 2531–2540.

Jørgensen, O. A., Hvingel, C., Møller, P. R., & Treble, M. A. (2005). Identification and mapping of bottom fish assemblages in Davis Strait and southern Baffin Bay. Canadian Journal of Fisheries and Aquatic Sciences, 62(8), 1833-1852. doi: 10.1139/f05-101

Jubinville, I., Lawler, E., Tattrie, S., Shackell, N.L., Flemming, J.M. & Worm, B. (2021) Distributions of threatened skates and commercial fisheries inform conservation hotspots. *Marine Ecology Progress Series*, *679*, pp.1–18.

Kneebone, J., Sulikowski, J., Knotek, R., McElroy, W.D., Gervelis, B., Curtis, T., Jurek, J. & Mandelman, J. (2020) Using conventional and pop-up satellite transmitting tags to assess the horizontal movements and habitat use of thorny skate (*Amblyraja radiata*) in the Gulf of Maine. *ICES Journal of Marine Science,* **77,** 2790–2803.

Krueger, F. (2019) Trim galore: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files.

554 Kulka, D.W., Ellis, J., Anderson, B., Cotton, C.F., Derrick, D., Pacoureau, N. & Dulvy, N.K.
555 (2020) *Amblyraja radiata*. *The IUCN Red List of Threatened Species,* 2020, e.
556 T161542A124503504. https://dx.doi.org/10.2305/IUCN.UK.2020-
557 3.RLTS.T161542A124503504.en
558 Lanfear R, Frandsen PB, Wright AM, Senfeld T, & Calcott B. (2017) PartitionFinder 2: New
559 Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological
560 Phylogenetic Analyses. *Molecular Biology and Evolution,* **34**, 772–773.
561 Leigh, J.W. & Bryant, D. (2015) POPART: full-feature software for haplotype network
562 construction. *Methods in Ecology and Evolution,* **6,** 1110–1116.
563 Li, C., Hofreiter, M., Straube, N., Corrigan, S. & Naylor, G.J. (2013) Capturing protein-coding
564 genes across highly divergent species. *BioTechniques,* **54,** 321–326.
565 Lynghammar, A., Christiansen, J.S., Griffiths, A.M., Fevolden, S.E., Hop, H. & Bakken, T.
566 (2014) DNA barcoding of the northern Atlantic skates (Chondrichthyes, Rajiformes),
567 with remarks on the widely distributed starry ray. *Zoologica Scripta,* **43,** 485–495.
568 Lynghammar, A., Præbel, K., Bhat, S., Fevolden, S.E. & Christiansen, J.S. (2016) Widespread
569 physical mixing of starry ray from differentiated populations and life histories in the
570 North Atlantic. *Marine Ecology Progress Series,* **562,** 123–134.
571 Matzke, N.J. (2013). Probabilistic historical biogeography: New models for founder-event
572 speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
573 *Frontiers of Biogeography*, **5,** 242–248.
574 McPhie, R. & Campana, S. (2009) Reproductive characteristics and population decline of four
575 species of skate (Rajidae) off the eastern coast of Canada. *Journal of Fish Biology,* **75,**
576 223–246.

577 Mecklenburg, C. W., Lynghammar, A., Johannesen, E., Byrkjedal, I., Christiansen, J. S., Dolgov,
578 A. V., . . . Wienerroither, R. (2018) *Marine fishes of the Arctic region*. Conservation of
579 Arctic Flora and Fauna, Akureyri, Iceland., ISBN 978-9935-431-69-1. Available via
580 https://www.caff.is/marine/marine-monitoring-publications/marine-fishes-of-the-arctic-
581 region.
582 Meyer, M. & Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed
583 target capture and sequencing. *Cold Spring Harbor Protocols,* **2010,** pdb. prot5448.
584 Naylor, G. J., Caira, J. N., Jensen, K., Rosana, K. A., Straube, N., & Lakner, C. (2012).
585 Elasmobranch phylogeny: a mitochondrial estimate based on 595 species. *Biology of*
586 *sharks and their relatives*, *2*, 31-56.
587 New England Fishery Management Council (NEFMC) (2003) Final fishery management plan for
588 the Northeast Skate Complex. Newburyport, MA. 443 p.
589 National Marine Fisheries Service (NMFS) (2017) Status Review Report: Thorny Skate
590 (*Amblyraja radiata*). *Final Report to National Marine Fisheries Service, Office of*
591 *Protected Resources*, pp. 60.
592 Nye, J.A., Link, J.S., Hare, J.A. & Overholtz, W.J. (2009) Changing spatial distribution of fish
593 stocks in relation to climate and population size on the Northeast United States
594 continental shelf. *Marine Ecology Progress Series*, ***393*****,** 111–129.
595 Ostrow, D.G., Salomon, M.P., Clark, G., Kneebone, J., Sulikowski, J., & Tsang, P.C.W. (2008)
596 Population genetic patterns among phenotypically divergent thorny skate (*Amblyraja*
597 *radiata*) populations from the western Gulf of Maine. 24th Annual Meeting of the
598 American Elasmobranch Society, Montreal, Quebec, Canada. July 2008.

599 Rohland, N. & Reich, D. (2012) Cost-effective, high-throughput DNA sequencing libraries for
600 multiplexed target capture. *Genome Research,* **22,** 939–946.
601 Rubinoff, D., Cameron, S. & Will, K. (2006) A genomic perspective on the shortcomings of
602 mitochondrial DNA for "barcoding" identification. *Journal of Heredity,* **97,** 581–594.
603 Sguotti, C., Lynam, C.P., García-Carreras, B., Ellis, J.R. & Engelhard, G.H. (2016) Distribution
604 of skates and sharks in the North Sea: 112 years of change. *Global Change*
605 *Biology*, ***22***(8), 2729–2743.
606 Smith, S.A. & O'Meara, B.C. (2012) treePL: divergence time estimation using penalized
607 likelihood for large phylogenies. *Bioinformatics,* **28,** 2689–2690.
608 Sosebee, K.A. (2005) Maturity of skates in northeast United States waters. *Journal of Northwest*
609 *Atlantic Fisheries Science,* **35,** 141–153.
610 Stamatakis, A. (2006) RAxML–VI–HPC: Maximum Likelihood-based Phylogenetic Analyses
611 with Thousands of Taxa and Mixed Models. *Bioinformatics,* **22**, 2688–2690.
612 Stamatakis, A. (2014) RAxML version 8: A tool for phylogenetic analysis and post- analysis of
613 large phylogenies. *Bioinformatics,* **30**, 1312–1313.
614 Stamatakis, A., Hoover, P., & Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML
615 web-servers. *Systematic Biology* **57**, 758–771.
616 Stortini, C.H., Shackell, N.L., Tyedmers, P. & Beazley, K. (2015) Assessing marine species
617 vulnerability to projected warming on the Scotian Shelf, Canada. *ICES Journal of Marine*
618 *Science*, *72*(6), pp.1731–1743.
619 Sulikowski, J.A., Kneebone, J., Elzey, S., Jurek, J., Danley, P.D., Howell, W.H. & Tsang, P.C.
620 (2005) Age and growth estimates of the thorny skate (*Amblyraja radiata*) in the western
621 Gulf of Maine. *Fishery Bulletin*, **103,** 161–168.

622 Sulikowski, J.A., Kneebone, J., Elzey, S., Jurek, J., Howell, W.H. & Tsang, P.C.W. (2006) Using
623 the composite variables of reproductive morphology, histology and steroid hormones to
624 determine age and size at sexual maturity for the thorny skate *Amblyraja radiata* in the
625 western Gulf of Maine. *Journal of Fish Biology*, **69**(5), 1449–1465.
626 Swofford, D.L. (2002) PAUP: Phylogenetic Analysis Using Parsimony (and other methods),
627 Version 4.0b10. Sinauer Associates, Sunderland, MA.
628 Templeman, W. (1984) Migrations of thorny skate, *Raja radiata*, tagged in Newfoundland.
629 *Journal of Northwest Atlantic Fishery Science,* **5,** 55–63.
630 Templeman, W. (1987a) Differences in sexual maturity and related characteristics between
631 populations of thorny skate (*Raja radiata*) in the Northwest Atlantic. *Journal of*
632 *Northwest Atlantic Fisheries Science,* **7,** 155–167.
633 Templeman, W. (1987b) Length-weight relationships, morphometric characteristics and
634 thorniness of thorny skate (*Raja radiata*) from the Northwest Atlantic. *Journal of*
635 *Northwest Atlantic Fishery Science,* **7,** 89–98.
636 Templeton, A.R., Crandall, K.A. & Sing, C.F. (1992) A cladistic analysis of phenotypic
637 associations with haplotypes inferred from restriction endonuclease mapping and DNA
638 sequence data. III. Cladogram estimation. *Genetics,* **132,** 619–633.
639 Tsang, P., Sulikowski, J., & Danley, P. (2008) Sea Grant Project Termination Report. Age and
640 Size Difference in Sexually Mature Thorny Skates: Implications for Fisheries
641 Management in the Gulf of Maine. University of New Hampshire Sea Grant Program,
642 9pp.
643 Vihtakari, M. (2021) ggOceanMaps: plot data on oceanographic maps using 'ggplot2'. version
644 1.0.9. https://mikkovihtakari.github.io/ggOceanMaps.

Walker, P., Howlett, G., & Millner, R. (1997). Distribution, movement and stock structure of three ray species in the North Sea and eastern English Channel. *ICES Journal of Marine Science*, *54*(5), 797-808.

White, W.T., Corrigan, S., Yang, L., Henderson, A.C., Bazinet, A.L., Swofford, D.L. & Naylor, G.J. (2018) Phylogeny of the manta and devilrays (Chondrichthyes: Mobulidae), with an updated taxonomic arrangement for the family. *Zoological Journal of the Linnean Society,* **182,** 50–75.

Figure captions

Figure 1. Sampling locations for 527 individual thorny skate, *Amblyraja radiata*, for which whole mitogenomes were sequenced. Samples are color-coded by geographic region into which they were grouped by K-means clustering. Green shading denotes species distribution reported in Mecklenburg *et al.* (2018). Map is plotted in the Arctic Polar Stereographic coordinate system (EPSG:3995) using data associated with the R package *ggOceanMaps* (Vihtakari, 2021). Gulf of Maine (GoM) and Newfoundland (NL), and Northeast including South Greenland (SG), East Greenland (EG), West Iceland (WI), East Iceland (EI), North Sea (Nsea), Northern Norway (Nnor), and Arctic Circle (ArcC).

Figure 2. Haplotype network of the protein-coding mitogenome (N = 11,391 bp) for the 527 individual thorny skate, *Amblyraja radiata*, and 443 haplotypes analyzed in this study. Individuals are color-coded by the two major geographic groups: Northwest containing the Gulf of Maine (GoM) and Newfoundland (NL), and Northeast including South Greenland (SG), East

668 Greenland (EG), West Iceland (WI), East Iceland (EI), North Sea (Nsea), Northern Norway
669 (Nnor), and Arctic Circle (ArcC). The number of dashes along a branch indicated the number of
670 mutational steps separating the haplotypes. Black dots are extinct or unsampled intermediate
671 haplotypes. See Figure S1 for the same haplotype network color-coded by the nine geographic
672 regions.
673
674 Figure 3. Timetree built based on 443 unique protein-coding mitogenome sequences of thorny
675 skate, *Amblyraja radiata*, rooted with two outgroups (*Amblyraja doellojuradoi* and *Rajella*
676 *fyllae*; not shown; see Figure S3 for the full timetree). For major groupings, divergence times are
677 shown above branches, while bootstrap support values are shown below branches. Taxa and
678 clades in red indicate GoM and NL samples. The ancestral range reconstruction results from
679 BioGeoBEARS analyses (see Figure S4 for detailed results) suggested that there might have
680 been 28 historical East - West colonization events (indicated by blue stars) and four historical
681 West - East colonization events (indicated by red stars).
682
683 **Figure captions for supplementary figures**
684 Figure S1. Haplotype network of the protein-coding mitogenome (N = 11,391 bp) for the 527
685 individual thorny skate, *Amblyraja radiata*, analyzed in this study. Samples are color-coded by
686 the nine geographic populations, as follows: GoM: Gulf of Maine, NL: Newfoundland, SG:
687 South Greenland, EG: East Greenland, WI: West Iceland, EI: East Iceland, Nsea: North Sea,
688 Nnor, Northern Norway, ArcC: Arctic Circle. The number of dashes along a branch indicated the
689 number of mutational steps separating the haplotypes. Black circles are extinct or unsampled
690 intermediate haplotypes.

691

692 Figure S2. The best tree (*Ln*L = -38275.365145) resulting from Maximum Likelihood analysis of
693 443 unique protein-coding mitogenome sequences of thorny skates, *Amblyraja radiata*. Numbers
694 near nodes are bootstrap support values.

695

696 Figure S3. The full timetree built based on protein-coding mitogenome for 527 individual thorny
697 skate, *Amblyraja radiata*, analyzed by this study. Horizontal bars denote 95% Highest Posterior
698 Density (HPD) intervals of node ages, which were calculated from the 500 independent timetrees
699 built using TreePL. Samples from the West group (GoM/NL) are highlighted in red.

700

701 Figure S4. Ancestral range reconstruction for 527 individual thorny skate, *Amblyraja radiata*,
702 based on the DIVALIKE model in BioGeoBEARS. The MCC tree generated by our divergence
703 time analyses were used as input phylogenies. We defined nine biogeographic areas (A - I) that
704 correspond to the nine regions (GoM, NL, SG, EG, WI, EI, Nsea, Nnor, and ArcC), respectively.
705 The maximum number of ancestral areas allowed at each node was set as four.

706

707 Figure S5. Mapping size morph information for 527 individual thorny skate, *Amblyraja radiata,* on
708 the haplotype network (Figure 2). Samples were color-coded as "large", "small" or "NA" (not in Gulf of
709 Maine, or size information unknown). The number of dashes along a branch indicated the number of
710 mutational steps separating the haplotypes. Small black dots are extinct or unsampled intermediate
711 haplotypes.

Table 1. Molecular diversity measurements by group and geographic region for 527 individual thorny skate, *Amblyraja radiata*. N = sample size; H = number of haplotypes; *h* = haplotypic diversity; π = nucleotide diversity. Bold values are significant at the $p < 0.001$ level.

Group and region	N	H	*h*	π	Fu's *Fs*
Northwest	351	284	0.9985 +/- 0.0004	0.0054 +/- 0.0026	**-23.7973**
- Gulf of Maine (GoM)	293	229	0.9979 +/- 0.0005	0.0052 +/- 0.0025	**-23.8423**
- Newfoundland (NL)	58	55	0.9982 +/- 0.0036	0.0061 +/- 0.0030	-11.2610
Northeast	176	159	0.9986 +/- 0.0008	0.0049 +/- 0.0024	**-24.0521**
- Arctic Circle (ArcC)	46	41	0.9942 +/- 0.0062	0.0047 +/- 0.0023	-5.4025
- East Greenland (EG)	3	3	1.0000 +/- 0.2722	0.0037 +/- 0.0028	2.6233
- East Iceland (EI)	38	36	0.9972 +/- 0.0068	0.0053 +/- 0.0026	-5.4007
- Northern Norway (Nnor)	9	9	1.0000 +/- 0.0524	0.0051 +/- 0.0028	0.2140
- North Sea (Nsea)	11	9	0.9455 +/- 0.0659	0.0048 +/- 0.0026	2.9506
- South Greenland (SG)	11	11	1.0000 +/- 0.0388	0.0056 +/- 0.0029	-0.2436
- West Iceland (WI)	58	54	0.9970 +/- 0.0041	0.0046 +/- 0.0022	-13.3252
All samples	527	443	0.9992 +/- 0.0002	0.0063 +/- 0.0030	-

Table 2: Results of one-way hierarchical analysis of molecular variance (AMOVA) assessing geographic clusters of thorny skate, *Amblyraja radiata*, within a Northwest/Northeast structuring. Percent of total variance explained is presented in parentheses. df = degrees of freedom, SS = sum of squares

Source of variation	df	SS	Variance Component
Among groups	1	3025.254	12.290 Va (28.69%)
Among populations within groups	7	442.749	1.019 Vb (2.38%)
Within populations	518	15294.855	29.527 Vc (68.93%)
Total	526	18762.858	42.836

Table 3 – Pairwise F_{ST} values between regions for which at least 10 samples of thorny skate, *Amblyraja radiata*, were analyzed. Bold values are significant at the $p < 0.01$ level. Gulf of Maine (GoM), Newfoundland (NL), South Greenland (SG), East Greenland (EG), West Iceland (WI), East Iceland (EI), North Sea (Nsea), Northern Norway (Nnor), and Arctic Circle (ArcC)

	GoM	NL	Nsea	SG	WI	EI
NL	**0.0556**					
Nsea	**0.3584**	**0.1880**				
SG	**0.3139**	0.1225	0.0416			

WI	**0.3358**	**0.1654**	0.0376	0.0074		
EI	**0.3084**	**0.1301**	0.0272	-0.0163	0.0011	
ArcC	**0.3454**	**0.1701**	0.0785	0.0243	0.0057	0.0091

Table 4: Results of a two-way analysis of molecular variance (AMOVA) examining the effect of sex × region across all thorny skate, *Amblyraja radiata*, samples with both sex and region information available (N = 482). Statistical significance was accepted at the $p < 0.01$ level. df = degrees of freedom, SS = sum of squares

	df	SS	R^2	F	p-value
Sex	1	0.0000304	0.00259	1.7015	0.156
Region	8	0.0032881	0.28013	22.9702	0.001
Sex × Region	8	0.0001165	0.00993	0.8141	0.671
Residuals	464	0.0083026	0.70734		
Total	481	0.0117377	1		

Table 5: Results of a two-way analysis of molecular variance (AMOVA) examining the relationship between sex × size morph for 148 specimens of thorny skate, *Amblyraja radiata*, collected in the Gulf of Maine for which sex and maturity status was available. Statistical significance was accepted at the $p < 0.01$ level. df = degrees of freedom, SS = sum of squares

	df	SS	R^2	F	p-value
Sex	1	0.00000955	0.00339	0.4922	0.589
Morph	1	0.00000513	0.00182	0.2645	0.823
Sex × Morph	1	0.00000413	0.00147	0.2127	0.899
Residuals	144	0.00279315	0.99331		
Total	147	0.00281195	1		

Table S1: Sample metadata and GenBank numbers for 527 individual thorny skate, *Amblyraja radiata*, analyzed by this study. GoM: Gulf of Maine, NL: Newfoundland, SG: South Greenland, EG: East Greenland, WI: West Iceland, EI: East Iceland, Nsea: North Sea, Nnor, Northern Norway, ArcC: Arctic Circle.