ORIGINAL PAPER

Population dynamics of *Arctica islandica* off Long Island (USA): an analysis of sex-based demographics and regional comparisons

Kathleen M. Hemeon[1,3] · Eric N. Powell[1] · Sara M. Pace[1] · Roger Mann[2] · Theresa E. Redmond[2]

Abstract

The boreal bivalve *Arctica islandica* is an important fishery in the United States (US), yet very little is known about the resiliency of this species to fishing activity due to limited understanding of localized population demographics. Demographics including age frequency, recruitment patterns, mortality rates, and sexual dimorphism were evaluated for a population sampled off Long Island (LI, 40.09658°N 73.01057°W) and compared with samples from Georges Bank (GB, 40.72767°N, 67.79850°W) collected in 2015 and 2017, where GB was described in a previous study. This study supports evidence that this species is sexually dimorphic. Earlier assumptions of prolonged lapses in recruitment were not substantiated for either the GB or LI populations; yearly cohorts were observed for the past century, and both populations presented recruitment pulses in approximately 8-y periods. Estimated ages from this study are older than previously reported for the US Mid-Atlantic with the oldest animal represented by a 310-year-old male collected from LI. Simulated total mortality was higher at GB than LI, and higher for GB females than GB males, with simulated mean longevity estimates greater at LI than GB. The population sex ratio at GB was 1:1.1 (female:male), whereas the LI ratio was 1:1.4 and relatively deficient in large females. Recruitment into the populations occurs routinely with substantial hiatuses being rare and substantive year classes occurring at least decadally with lesser, but contributing, recruitment in most years in between. Routine recruitment may insulate this species from risks posed by overfishing to an extent not typical for other long-lived species.

Keywords *Arctica islandica* · Dimorphism · Mortality · Longevity · Recruitment · Age–length keys · Ocean quahog

Introduction

The ocean quahog *Arctica islandica* is the longest lived abundant bivalve on Earth. Individuals on the deep continental shelf of the Mid-Atlantic (US) can survive for centuries, and when found in the colder, boreal waters of Iceland, ages over 500 years can be reached. The ocean quahog is a commercially harvested clam in the US Mid-Atlantic, yet very little is known regarding local recruitment frequency and

regional population dynamics of the US stock. A federally managed species, *A. islandica*, is managed as a single unit that combines two area-specific assessment models to create a single harvest quota (excluding the limited Gulf of Maine fishery). One assessment model analyzes the northern portion of the stock at Georges Bank (GB) (federally referenced as GBK-North) and the second model analyzes the southern portion of the stock from Virginia in the south to the western edge of the Great South Channel to the northeast (federally referenced as Southern Virginia/Southern New England [SVA/SNE]-South) (Fig. 1). Georges Bank is modeled as a separate management area due to its distinctive oceanographic setting, limited contemporary harvests, and more restricted survey data (NEFSC 2020). Unlike many other federally managed species, the *A. islandica* fishery quota is derived using only length-based assessments since age compositions are notoriously difficult to assemble for this species (NEFSC 2020). Even the traditional use of growth curves or age–length keys (ALK) to estimate age at size are currently not employed due to the extreme variability of age

✉ Kathleen M. Hemeon
   kathleen_hemeon@fws.gov

1   Gulf Coast Research Laboratory, University of Southern Mississippi, Ocean Springs, MS 39564, USA

2   Virginia Institute of Marine Science, College of William and Mary, Gloucester Point, VA 23062, USA

3   Abernathy Fish Technology Center, U.S. Fish and Wildlife Service, Longview, WA 98632, USA

🙋 Springer
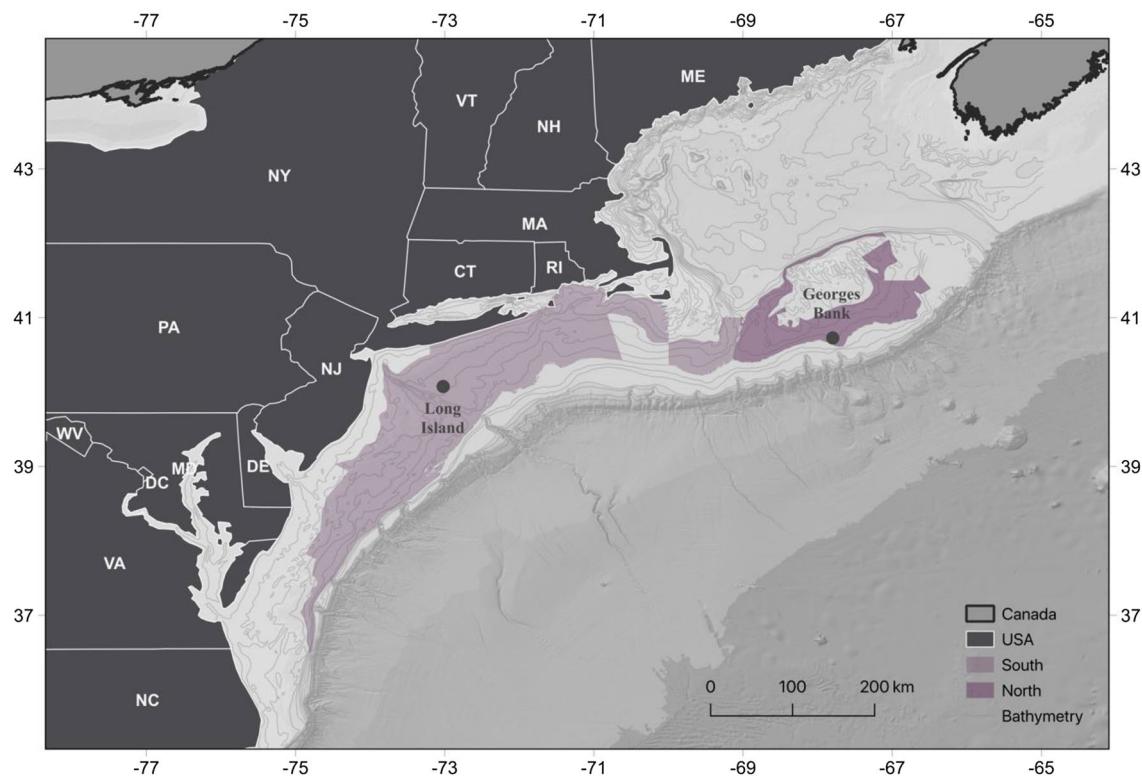
Fig. 1 Regional map of Mid-Atlantic *Arctica islandica* fishery management areas (polygons). The southern Management Area (South, SVA-SNE) is represented by the left, lighter polygon, and the Northern Management Area (North, GBK) is repre- sented by the right, darker polygon. The Long Island site is in management strata 4Q, and the Georges Bank site is in management strata 9Q

at size (Pace et al. 2017a,b; Hemeon et al. 2021a), although Hemeon et al. (2021a) did develop defensible ALKs for GB using large age–length datasets.

Studies spanning the North Atlantic demonstrate distinct regional and sex-specific growth dynamics for this species (Ropes et al. 1984b; Thorarinsdóttir and Steingrímsson 2000; Ridgway et al. 2012; Pace et al. 2018; Poitevin et al. 2019; Hemeon et al. 2021a), results that also suggest ALKs used to estimate age at length and time to maturity are also regionally and sexually explicit. Moreover, it is unclear at what spatial scale age and growth dynamics might diverge because of local oceanographic processes. A detailed age–length analysis was conducted for a site on GB to decipher population demographics specific to the Mid-Atlantic using a quasi-virgin *A. islandica* population (Hemeon et al. 2021a). Hemeon et al. (2021a) found that longevity and mortality estimates were higher than previously documented for *A. islandica* populations along the US continental shelf, that sex-ratios at size were statistically different, and that male and female *A. islandica* growth dynamics were dimorphic, warranting separate sex-based analyses (e.g., ALKs, mortality, longevity). Georges Bank is a relatively unfished population and

is located at a considerable distance offshore compared to other management areas to the west. The fishery off Long Island (LI) produces the largest fraction of total US landings, but a comprehensive population dynamics study has not been conducted for the LI population. Enhanced knowledge of *A. islandica* sexual dimorphism, sex ratios, mortality rates, and applications of age–length keys would vastly improve not only sustainable fisheries management, but also continental shelf ecological research and future aging studies for paleo-oceanographic reconstruction.

The objectives of this study were to describe sex-specific population demographics of *A. islandica* collected from the LI portion of the US stock (within federal management area SVA/SNE-South) and to perform a comparative analysis with the population derived from GB (federal management area GBK-North, Hemeon et al. 2021a). Population demographics analyzed included length frequencies, ALKs, age frequencies, sex ratios at size, mortality, and longevity. As the dominant harvest region, this study addressed the imperative that population-scale demographics be available to fishery managers, and that differences between intra-stock populations be identified.

Materials and methods

Sample collection

Two types of field samples were collected from LI: a sample to estimate length frequency (sampled in 2015) and a shucked sample to obtain sex ratios and animals for subsequent aging (sampled in 2017). Both samples were collected from approximately 40.09658°N 73.01057°W at a depth of 48 m. This study was designed to evaluate the demographics of *Arctica islandica* available to the fishery; therefore, only animals commonly selective to commercial gear were retained for analysis and, at LI, this included clams greater than 60 mm in shell length. The length sample was collected in 2015 with a commercial hydraulic clam dredge used by both the fishery and federal surveys (Hennen et al. 2016), and only shell lengths were measured. Formerly, this length sample was used alongside a 2015 age analysis by Pace et al. (2017a,b), and the length data were reutilized for the present study.

Results from Pace et al. (2017a,b) suggested that a more intensive age-at-length analysis using a larger age data-set was necessary for LI to fully explain local population dynamics. Therefore, in 2017, the second sample from this same location was collected with different sampling gear to collect a larger sample to be used for a new age analysis. The 2017 sample, hereafter termed the shucked sample, was collected using a Dameron-Kubiak (DK) dredge that can target smaller *A. islandica* and allowed more animals to be retained from the smaller length classes (herein referred to as size classes) of the fishery (60–80 mm) (Hennen et al. 2016). To meet desired age sample-size specifications (~ 100 animals per 5-mm size class), multiple 5-min tows were required. The shucked sample included all animals collected in the DK dredge. To assure an unbiased sample of the population for further analysis, all animals were measured for shell length, tissue shucked from the shell, sex determined by gonadal smear slide, and shell valves retained in dry storage for future analysis (Pace et al. 2017a; Hemeon et al. 2021a).

Sample preparation

The 2017 shucked sample was subsampled and approximately 100 animals per 5-mm size class were visually aged (herein referred to as the age sample). If available, each size class included 50% randomly selected females and 50% randomly selected males. All samples were aged for rare size classes with less than 100 animals. Shells to be aged were cross-sectioned using a tile saw, ground and polished to a reflective finish, and the hinge plate imaged by cellSens software and a high-resolution Olympus microscope (see Pace et al. 2017a; Hemeon et al. 2021b). ImageJ software (ObjectJ plugin) was used for image annotation to allow precise aging.

Error assessment

Growth lines in *A. islandica* are annually resolved (Jones 1980; Murawski et al. 1982, Ropes et al. 1984a, b; Weidman et al. 1994; Schöne et al. 2005; Mette et al. 2016; Reynolds et al. 2016), but great care must be given to consistently age the annuli and to omit secondary (i.e., subannual) growth lines that may be produced at young ages when growth rates are high. To standardize age results, a primary and secondary age reader each aged a random 20% subsample of the 2017 age sample and age-reader error metrics were evaluated using a threefold error analysis of precision, bias, and extreme error frequency (Hemeon et al. 2021b). This error protocol ensured that precision between two age readers was within acceptable thresholds, determined that no aging bias occurred, and permitted the assessment of the frequency of extreme deviations of precision (i.e., extreme errors). The error criteria proposed by Hemeon et al. (2021b) were targeted and included a median coefficient of variation less than 7%, an Evans–Hoenig bias P value greater than 0.05, and an extreme error frequency of less than 10% of the sample. Once error was constrained, the primary age reader aged the entire age sample.

Length frequency

The 2015 length sample was adjusted for dredge selectivity to better reflect the true length frequency of the population (see Table 15 in NEFSC 2017). The population length frequency was then divided into male and female length frequencies by use of sex proportion at size (Table 1). In other words, the length sample was multiplied by *A. islandica* length-specific dredge selectivity factors, and the selectivity-adjusted length frequency was divided into male and female groups using sex proportions within 1-mm size classes derived from the shucked sample (Table 2).

Male and female length frequencies were compared using the Kolmogorov–Smirnov (KS) (Conover 1980), Wald–Wolfowitz Runs (Runs) (Conover 1980), and Anderson–Darling (AD) (Pettitt 1976; Engmann and Cousineau 2011) statistical tests with *A. islandica*-specific modifications listed in Hemeon et al. (2021a). Mean male and female lengths were also compared using the Mann–Whitney *U* test in the R base statistics function (R Core Team 2020). The population length frequency was not independent from the sex-specific length frequencies as the population length frequency was divided into the male and female length frequencies using sex-ratio proportions at size.

Table 1 Long Island length frequency

Size class (mm)	Population	Female	Male
60	14	7	7
65	52	6	46
70	219	40	179
75	408	86	322
80	661	234	427
85	570	232	338
90	570	310	260
95	282	182	100
100	100	82	18
105	28	25	3
110	1	1	0
Total	2905	1205	1700

The length sample was adjusted for dredge selectivity to create the population length frequency, and sex proportions by size were applied to the selectivity-adjusted length frequency to create female and male length frequencies

Age frequency

The original age sample was expanded to include more shells than required (~ 100 per size class), as additional shells were imaged to replace those omitted due to poor image quality and all available ages were retained for such a data-poor species. Therefore, 904 ages were available to create the 2017 ALK binned into 5-mm size classes. Separate ALKs were created for the population, male, and female groups. Corresponding 2015 length frequencies were applied to the 2017 ALKs to create 2015 age frequencies for all three groups. Ceiling rounds were applied to the age frequency results to prevent the elimination of ages represented by fractional animals at rare ages.

Population, male, and female age frequencies were compared using the KS, Runs, and AD tests using an $\alpha = 0.05$ significance level. Means of population, male, and female age frequencies were also compared using a type III ranked one-way ANOVA (R Core Team 2020) with Tukey post hoc analyses.

Age–length key validation

For each sex, ALK reliability was analyzed using 50 Monte Carlo simulations (herein referred to as base simulations) to re-select age data and create 50 new ALKs and 50 new age frequencies (Fig. 2). Simulated datasets were obtained by choosing with replacement the same number of animals as in each original 5-mm size class using the Knuth Ran1 and Ran3 random number generators alternately, with the generator reinitialized by a new seed number for each simulation (Press et al. 1989). Age frequencies were then compared with the original group-specific (e.g., population, male, female) age frequency using the KS, AD, and Runs tests. The probability of a significantly different test across simulations was reported.

To test whether the population, male, and female ALKs were in fact unique, an additional 50 simulations (herein referred to as "substituted-group simulations") were completed to develop 50 new ALKs, but the substituted-group length frequency was applied to the ALKs to create substituted-group age frequencies (Fig. 2). The probability of a significantly different test from the original group-specific age frequency across simulations was reported. If the ALKs are effectively the same and group-specific ALKs are not required, the probability of a significant test across all 50 substituted simulations would be similar to the probability of significant tests for the base simulations using binomial analysis (Table 3; Hemeon et al. 2021a).

To determine if LI and GB require different ALKs, group-specific simulations were completed in a similar fashion as the substituted simulations, where 50 simulations (herein referred to as "substituted-site simulations") were completed to develop 50 new site-specific ALKs, but the substituted-site, group-specific length frequency was applied to the ALKs to create substituted-site, group-specific age frequencies. Sex-specific ALK reliability for the GB population was evaluated by Hemeon et al. (2021a), but GB ALKs were also analyzed in this study for region-scale ALK verification. The probability of significantly different tests across simulations was reported. If the ALKs are effectively the same and site-specific ALKs are not required, the probability of a significant test across all 50 substituted simulations would be similar to the probability of significant tests for the base simulations for each site using binomial analysis (Table 3, Hemeon et al. 2021a).

Mortality and longevity

Due to the extremely rich age composition dataset, and the use of age-frequency-derived *A. islandica* mortality estimates from other regions (e.g., Ridgway et al. 2012), mortality was first estimated by linear regression of the age frequency (i.e., Ricker method; Ricker 1975). Age-frequency data were only used in this analysis for the peak and right-hand portions of the frequency distribution after grouping into 10-y age classes to minimize year-to-year variability. This technique prevents smaller animals not fully selected by the dredge from affecting the mortality estimate. For comparison, an alternative estimate that does not demand expansive age compositions was used, namely the Hoenig$_{nls}$ function that only requires the maximum age of the population (Eq. 2.1) (see Table 4; Then et al. 2015)

$$M_{est} = 4.899 t_{max}^{-0.916} \tag{2.1}$$

Table 2 Long Island sex proportions at size. All sex-determined samples from Long Island (i.e., the shucked sample) were divided into 1-mm size classes and female and male proportions (P) were calculated from the number (N) of females and males per size class

Length (mm)	P_{Female}	P_{Male}	N_{Female}	N_{Male}
61	0.50	0.50	1	1
65	0.00	1.00	0	1
66	0.00	1.00	0	2
67	0.00	1.00	0	5
68	0.25	0.75	2	6
69	0.20	0.80	2	8
70	0.19	0.81	3	13
71	0.07	0.93	1	14
72	0.18	0.82	2	9
73	0.32	0.68	8	17
74	0.12	0.88	2	15
75	0.10	0.90	1	9
76	0.24	0.76	5	16
77	0.21	0.79	5	19
78	0.29	0.71	8	20
79	0.24	0.76	8	25
80	0.38	0.63	12	20
81	0.26	0.74	9	26
82	0.42	0.58	18	25
83	0.36	0.64	16	28
84	0.36	0.64	14	25
85	0.33	0.67	16	33
86	0.30	0.70	10	23
87	0.34	0.66	14	27
88	0.56	0.44	25	20
89	0.61	0.39	31	20
90	0.51	0.49	28	27
91	0.53	0.47	31	28
92	0.57	0.43	39	29
93	0.56	0.44	35	27
94	0.58	0.42	43	31
95	0.57	0.43	44	33
96	0.58	0.42	49	35
97	0.67	0.33	54	27
98	0.75	0.25	66	22
99	0.81	0.19	54	13

Table 2 (continued)

Length (mm)	P_{Female}	P_{Male}	N_{Female}	N_{Male}
100	0.78	0.22	54	15
101	0.76	0.24	50	16
102	0.90	0.10	62	7
103	0.82	0.18	40	9
104	0.86	0.14	43	7
105	0.81	0.19	25	6
106	1.00	0.00	36	0
107	0.88	0.12	15	2
108	1.00	0.00	11	0
109	0.92	0.08	11	1
110	0.88	0.13	7	1

where M_{est} represents natural mortality and t_{max} represents maximum age. To understand variability in mortality and longevity estimates, the 50 simulated age frequencies used to evaluate ALK differences were also used to calculate 50 simulated mortality and longevity estimates per sex and per site by the Ricker (1975) method. The mean longevity estimate obtained from these simulations was used as the t_{max} value in the formula of Then et al. (2015).

Sex ratios

Sex ratios of the shucked sample were evaluated by 5-mm size classes using an exact binomial test for size classes where $n < 30$, and an approximate binomial test for size classes where $n > 30$ (R Core Team 2020). A population sex ratio was calculated from the length frequency after the length frequency was adjusted for selectivity and divided by sex. Sex ratios were also compared between regions using a two-sample binomial test under the same test conditions listed previously.

Results

Error assessment

A total of 158 samples were randomly selected from the LI dataset for an age-reader error assessment between two experienced *Arctica islandica* agers. This exercise ensured the primary age reader was aging consistently and to acceptable standards. Three rounds of error analysis (whereby a new 20% random subset was selected for each round) were required to meet species-specific error thresholds for precision and bias set forth in Hemeon et al. (2021b). Precision, as measured by median coefficient of variation (CV) across all samples, was 6% for the total population, 5% for females, and 6% for males. The precision target of < 7% was achieved and variability of age estimates between age readers was low. Median CV was used instead of the mean due to the extreme number of age classes (greater than 300 y) and high skewness of both age and error data. Evans–Hoenig bias results were non-significant ($P > 0.05$) for population (Chi-square test, $\chi^2 = 37.48$, $P = 0.11$), female (Chi-square test, $\chi^2 = 27.144$, $P = 0.30$, and male (Chi-square test, $\chi^2 = 30.68$, $P = 0.16$) samples, an indication that age readers were consistent with annuli detection and the non-bias error target was also achieved.

Error frequency (i.e., frequency of extreme error) was also evaluated for individual samples with CV greater than 10%, and error frequency was greater for LI than for GB as described in Hemeon et al. (2021b) and LI error frequency was 20% for the population, 13% for females, and 30% for males. The target extreme error frequency was set at 10% of
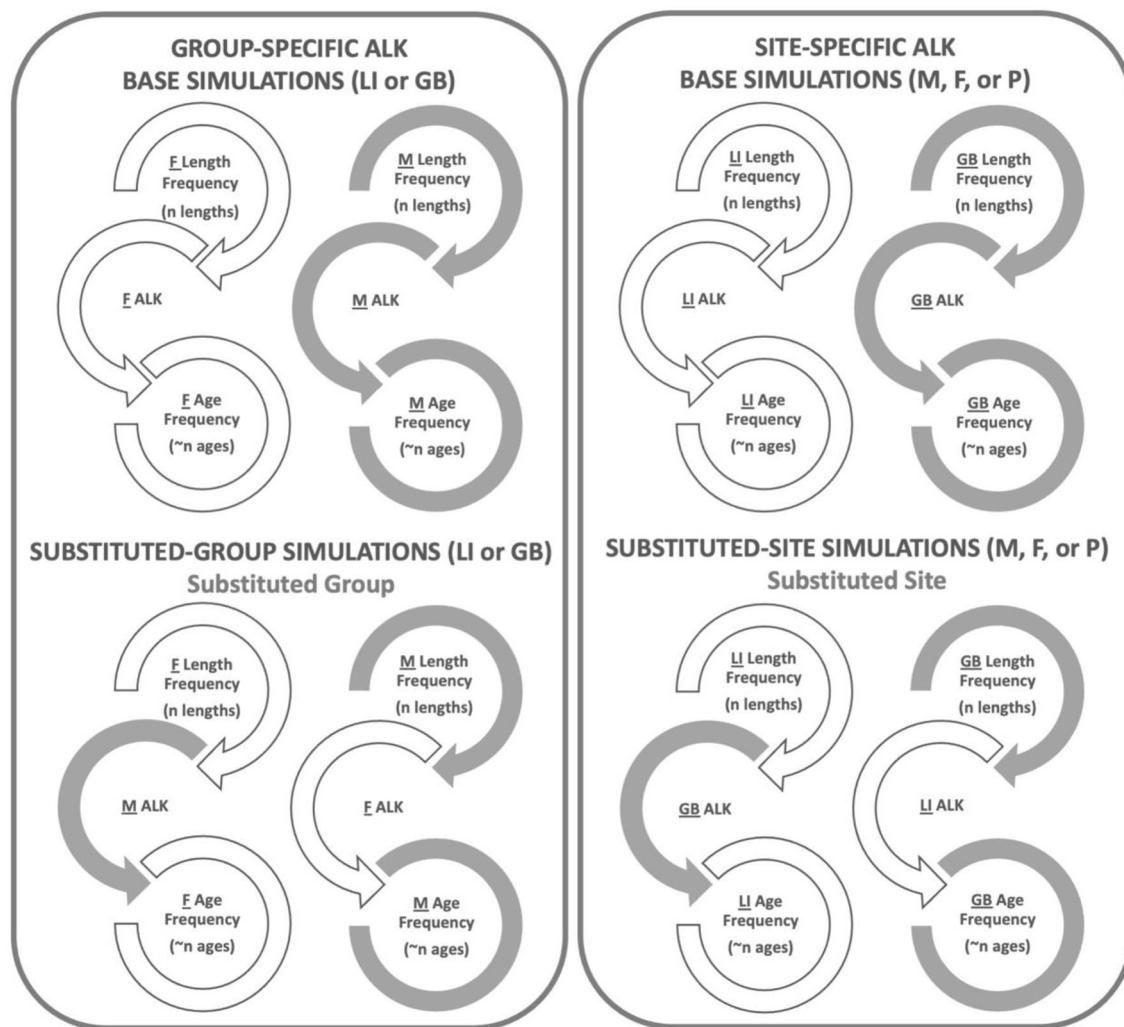
Fig. 2 Age–length key schematic. Age–length keys (ALK) were validated with base simulations for group-specific (population [P], female [F], male [M]) and site-specific (Long Island [LI], Georges Bank [GB]) ALKs. Substituted simulations used a substituted-group or substituted-site ALK to test differences in simulated age frequencies from true age frequencies

the sample. Samples collected from LI are more challenging to age due to the higher occurrence of suspected intra-annual growth lines that need to be interpreted by age readers, and thus resulted in more extreme cases of error between samples (particularly male samples). Due to high precision (CV) and no age-reader bias, error was deemed acceptable with the caveat that LI has higher error for less than 30% of the sample when compared to GB, and that males are aged with greater error than females at both sites.

Length frequency

The LI length frequency sample was adjusted for dredge selectivity (see Table 15 in NEFSC 2017) and divided into male and female datasets by sex proportion at size from the shucked sample (Table 2) to better reflect the true population. The adjusted length frequency included 1205 female and 1700 male shell lengths for clams available to the commercial (and federal survey) dredge (Table 1). Female median shell length was 89 mm (61–111 mm) with a mean length also of 89 mm ($\pm\,SD = 8$ mm, $n = 1205$). Male median shell length was 83 mm (61–107 mm) with a mean length also of 83 mm ($\pm\,SD = 8$ mm, $n = 1700$) (Fig. 3A). A Mann–Whitney–Wilcoxon test determined a significant difference between the mean female and male lengths ($W = 1.43\mathrm{e} + 6$, $P < 2.2\mathrm{e}\text{-}16$) (Fig. 3A), where female lengths are offset to larger sizes than males (Fig. 3B). The distribution statistics (Table 5) comparing LI male and female length frequencies were significant for the KS and Runs tests, but non-significant for the AD test. Results indicated that the tails of the distributions were similar, but the modal

Table 3 Georges Bank (GB) and Long Island (LI) age–length key validation

Analysis	True length frequency	ALK	KS	Runs	AD
Long Island	Female	Female	0.00	0.28	1.00
		Male	*1.00*	*0.50*	*1.00*
	Male	Male	0.00	0.26	1.00
		Female	*1.00*	*0.62*	*1.00*
	Population	Population	0.00	0.16	0.96
		Female	*0.84*	*0.68*	*1.00*
		Male	*1.00*	*1.00*	*1.00*
Georges Bank	Female[a]	Female	0.02	0.26	0.96
		Male	*1.00*	*0.26*	*1.00*
	Male[a]	Male	0.02	0.24	1.00
		Female	*1.00*	*0.32*	*0.82*
	Population	Population	0.00	0.40	0.78
		Female	*0.98*	*0.12*	*1.00*
		Male	*0.58*	*0.20*	*1.00*
Region	LI population	*GB Population*	*1.00*	*1.00*	*1.00*
	LI female	*GB Female*	*1.00*	*1.00*	*1.00*
	LI male	*GB Male*	*1.00*	*1.00*	*1.00*
	GB population	*LI Population*	*1.00*	*1.00*	*1.00*
	GB female	*LI Female*	*1.00*	*1.00*	*0.92*
	GB male	*LI Male*	*1.00*	*0.94*	*0.38*

Reliability and redundancy verification for site, sex, and regional age–length keys (ALK). A "True" designation indicates the length-frequency data applied to base and substituted ALKs, and the age frequency tested against simulation results from the Kolmogorov–Smirnov (KS), Runs, and Anderson–Darling (AD) tests. Grey highlighted cells signified when a substituted (italicized text) ALK produced simulations that were significantly different from the base simulations (plain text, bold) using a one-sample binomial test ($P < 0.05$). Expected proportions are those of the base simulations. Georges Bank analysis results were taken from Hemeon et al. (2021a). Regional analysis compares female to female, and male to male ALKs when the opposite-site same-sex ALK is substituted

[a]Data from Hemeon et al. (2021a)

Table 4 Mean mortality and longevity estimates with standard deviations (SD) from 50 simulated age frequencies using the Ricker (1975) method and associated linear regression fits (Adjusted R^2)

Site	Group	Mortality		Longevity		Adjusted R^2		Then et al. M_{est} (y^{-1})
		$\overline{M}$(y^{-1})	SD (y^{-1})	$\overline{L}$(y)	SD (y)	$\overline{R^2}$	SD	
Georges Bank	Female	0.042	3.17e-3	240	7.90	0.88	4.34e-2	0.032
	Male	0.035	2.91e-3	267	10.7	0.86	3.79e-2	0.029
	Population	0.042	2.22e-3	251	64.2	0.94	2.32e-2	0.031
Long Island	Female	0.020	1.69e-3	328	15.5	0.75	5.77e-2	0.024
	Male	0.018	2.45e-3	357	27.4	0.67	9.99e-2	0.022
	Population	0.020	1.90e-3	367	21.4	0.74	6.22e-2	0.022

Then et al. (2015) M_{est} (see Eq. 2.1) calculated using the Ricker $\overline{M}$ estimate

sections diverged and were offset (see conditions of the Runs test) (Fig. 3B).

Length frequencies at GB were consistently offset to larger sizes when compared to LI (Fig. 4). This trend was most apparent in the sex-specific length frequencies, where LI females were offset to smaller sizes than GB females (Fig. 4B) and the respective means were significantly different (Mann–Whitney–Wilcoxon test, $W = 1.33e + 6$, $P < 2.2e-16$). Long Island males were also offset to smaller sizes than GB males (Fig. 4C) (Mann–Whitney–Wilcoxon test, $W = 2.25e + 6$, $P < 2.2e-16$) (see Hemeon et al. 2021a for GB data). Regional length distributions were also evaluated, and tails were similar between sexes at the two sites, but the modal section of the distributions was different and both GB female and male length frequencies were offset to larger sizes (Table 5; Fig. 4).

Age–length data

The age sample included ages and lengths for 904 samples, including 448 female samples and 456 male samples. A target of 100 animals aged per 5-mm size class was exceeded for size classes 80–95 mm due to the availability of images (Table 6). The mean female age was 119 y ($\pm$ SD = 58 y, $n = 448$), and the mean male age was 107 y ($\pm$ SD = 58 y, $n = 456$). Size class 105 mm had the largest range of ages that spanned 249 y with a mean age of 181 y ($\pm$ SD = 46 y, $n = 86$) (Table 6) and the dataset ranged in ages from 17 to 310 y (median = 96 y) and lengths from 51 to 114 mm (median = 90 mm) (Fig. 5A). Only 0.1% of the aged sample was born in the past 20 y, and only 3% of the aged sample was born in the past 30 y. Limited numbers at young age is an artifact of the restriction of aging to individuals from larger size classes.

Age compositions described by sex and size class at LI were significantly different only for the 90-mm ($P = 6.70e-10$) size class using Tukey post hoc analysis, where females are younger than males within that size class (type III ranked two-way ANOVA, $F(1,10) = 2.35$, $P = 0.01$) (Fig. 5). When aged data are compared between LI and GB, females are statistically different regionally in the 95-mm ($P = 3.00e-4$) and 105-mm ($P = 2.00e-4$) size classes and Georges Bank females are younger than LI females within those size classes (type III ranked two-way ANOVA, $F(1,11) = 4.89$, $P = 6.68e-6$) (Fig. 5C). Males are statistically different regionally in the 70-mm ($P = 4.45e-05$), 75-mm ($P = 0.04$), 80-mm ($P = 2.00e-3$), and 90-mm ($P = 6.28e-8$) size classes in that GB males are younger than LI males within the 90-mm size class but older in the 70-mm, 75-mm, and 80-mm size classes (type III ranked two-way ANOVA, $F(1,11) = 14.73$, $P < 2e-16$) (Fig. 5D). The significantly older GB males in the 70-mm size class compared to the LI males

is likely not reliable due to the extremely small sample size of aged GB males in that size class ($n = 4$) (Hemeon et al. 2021a).

Age frequency

Age frequencies for population, female, and male groups were created independently with unique ALKs derived from the age sample (see Online Resources: Supplementary Material). The population age-frequency data ranged in age between 17 and 310 y (median = 84 y), female age-frequency data ranged between 17 and 272 y (median = 87 y), and male age-frequency data ranged between 21 and 310 y (median = 81 y) (Fig. 3). Type III ranked one-way ANOVA resulted in significant differences between population, female, and male groups (type III ranked one-way ANOVA, $F(2,6101) = 12.00$, $P = 6.32e$-6). Tukey post hoc analysis was significant between male and female ($P = 3.02e$-6), male and population ($P = 0.05$), and female and population ($P = 3.00e$-3) groups. Cumulative age frequencies for males and females do appear offset between 40 and 130 y, but otherwise track the other age frequency very closely (Fig. 3D).

Age-frequency distribution statistics (Table 5) identified a significant difference between LI male and female age frequencies only for the KS test, an indication that the frequencies are different in the modal portion of the distribution (Fig. 5) and likely near the median birth years of the two groups (~ 1927–1937) (Fig. 6). The population, female, and male age frequencies all present a large depression in the abundance of animals born during that time frame, but the male frequency is particularly deficient in animals born during 1920–1930. Notable reductions in effective recruitment (i.e., animals that survive to reach the fishery) also occurred between 1935 and 1945, and again in the 1970s. Younger animals (recent birth years) tend to be smaller and less available to the commercial dredge due to dredge size selectivity. An artificial drop in the frequency distribution of animals born in recent years is a consequence; however, isolated patterns can still be evaluated for the right-hand tail of the age-frequency distribution if the numbers are not compared to those on the left-hand tail due to gear selectivity effects. The population age frequency appeared to capture the modal section of both males and females since only the AD test is significantly different between the population age frequency and the two sex-specific age frequencies (Table 5).

Regional age frequencies are statistically different for all distribution tests for all groups (i.e., population, male, female), despite similar periods of depressed effective recruitment (i.e., animals born that recruit to the fishery) ~ 1920–1930 in both the LI and GB age frequencies (Fig. 6; Table 5; see also Hemeon et al. 2021a, Figs. 11, 12). Female age frequencies between LI and GB, and male age frequencies between LI and GB, were statistically significantly different for all three tests (Table 5). Therefore, differences in sex-specific age frequencies appear to exist both within and between populations of the Mid-Atlantic stock.

Age–length key validation

Simulations were used to evaluate ALKs for reliability, and versatility, across samples (e.g., group, site) (Table 3). The LI population ALK was reliable at replicating the modal section of the population age frequency (Fig. 6, Table 3), only producing an age-frequency distribution shift 16% of the time (see Runs results). Female and male LI ALKs were also reliable at the replication of the true sex-specific age-frequency modes but resulted in distribution shifts 26–28% of the time. The reliability of ALKs at both GB and LI are very similar, in that the modal section of the age-frequency distributions are reproducible greater than 98% of the time, the replicated age-frequency distributions may be offset in age between 16 and 28% of the time (although GB population ALK produced offsets 40% of the time), and the age-frequency distribution tails are generally not reproducible (GB population ALK is better than others at predicting the distribution tails) (Table 3). Thus, the long tail of old animals is least defined and remains poorly defined even with the large dataset used in this analysis to describe population age frequency.

Sex-specific ALKs are reliable, but not interchangeable within a region. In other words, a male or female ALK alone cannot replicate a true population age frequency. Likewise, ALKs are reliable at reproducing age-frequency distributions within a local population, but a single ALK from one location cannot be substituted for the other to represent a region (i.e., LI and GB combined).

Mortality and longevity

The age frequency for each LI group was evaluated for longevity using the Ricker method (Ricker 1975). Longevity was greatest for the population at 347 y ($R^2 = 0.91$, $F(1,21) = 226.5$, $P = 1.01e$-12), followed by females at 324 y ($R^2 = 0.90$, $F(1,19) = 182.2$, $P = 3.47e$-11), and finally males at 316 y ($R^2 = 0.89$, $F(1,20) = 160.1$, $P = 5.30e$-11) (Fig. 7), despite a male with the maximum observed LI age. Estimated longevity was greater than maximum observed age for all three groups, with female longevity being 19%, population 12%, and males 2% greater than the maximum observed age. Linear regression estimates of mortality were similar between population, females, and males at 0.022, 0.021, and 0.023 y^{-1}, respectively (multiple R^2 of 0.92, 0.91, and 0.89 respectively) (Fig. 7). The Ricker (1975) method was also
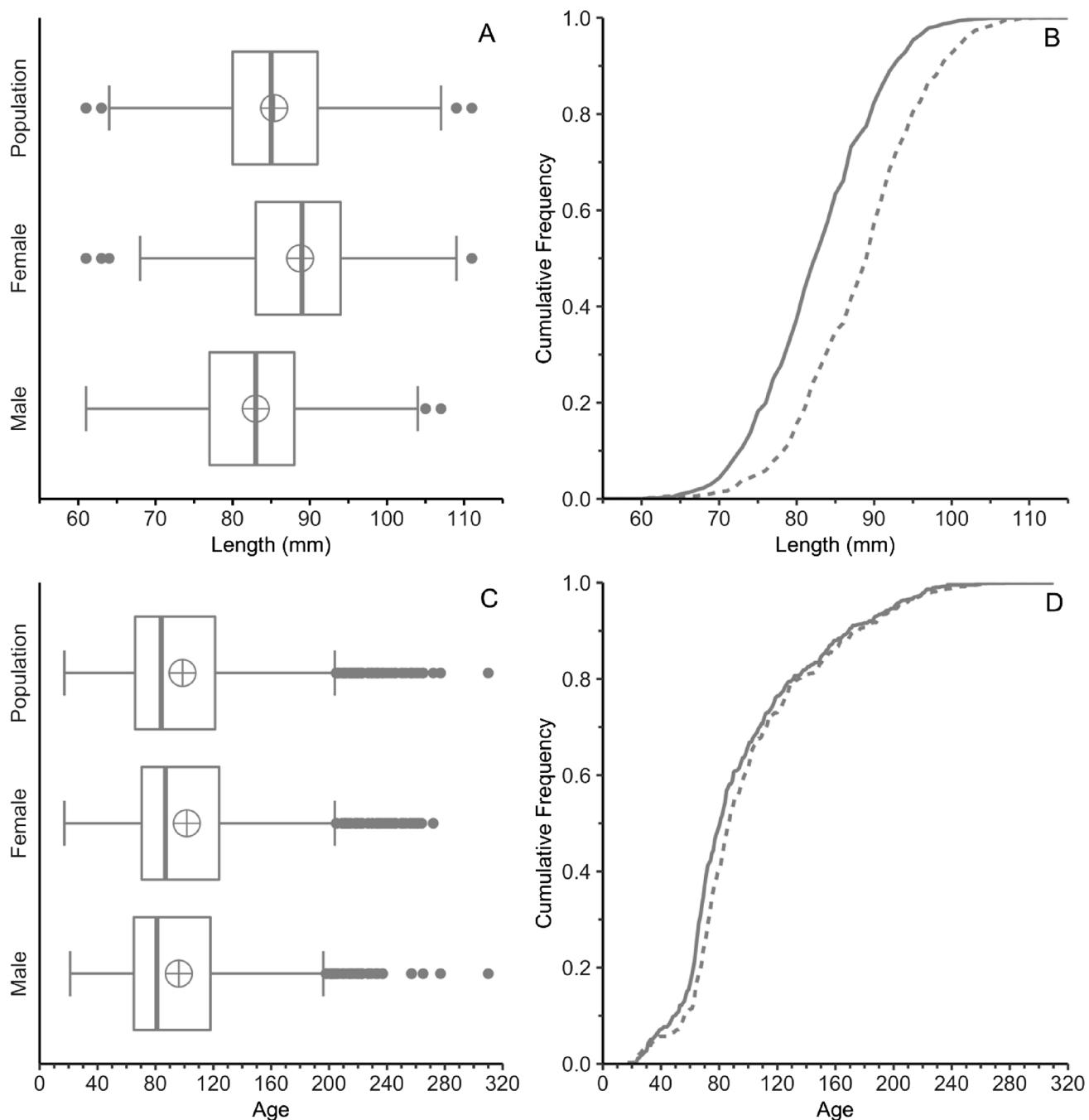
Fig. 3 Long Island length and age-frequency data summaries. **A** Length frequency data; **B** cumulative length frequencies of female (dashed) and male (solid) length data; **C** age-frequency data; **D** cumulative age frequencies of female (dashed) and male (solid) data. For boxplots, target represents mean, box represents the interquartile range (IQR) with 50th percentile bar (median), whiskers represent 1.5*IQR, and points are outliers

applied to the 50 simulated age frequencies developed for ALK validation (see Hemeon et al. 2021a for GB simulations). All simulated age frequencies produced significant linear regressions for each site by population, female, and male groups using this method. Estimates for mortality

using the Ricker (1975) method and these simulations supported the observed mortality estimates in that GB incurred higher rates of total mortality than LI, and that GB females had a higher natural mortality rate than GB males (Table 4; Fig. 8). Similarly, estimates of longevity from simulations were greater at LI than GB, and GB females represented

Table 5 Regional distribution statistics

Site	Dataset	Comparison	P value		
			KS	Runs	AD
Long Island	Length	M–F	< 0.01	< 0.01	ns
	Age	M–F	< 0.01	ns	ns
		P–F	ns	ns	< 0.01
		P–M	ns	ns	< 0.01
Region	Length	M–M	< 0.01	< 0.01	ns
		F–F	< 0.01	< 0.01	ns
	Age	M–M	< 0.01	< 0.01	< 0.01
		F–F	< 0.01	< 0.01	< 0.01
		P–P	< 0.01	< 0.01	< 0.01

Statistics for length- and age-frequency distributions between males and females, between population and sex-specific age-frequency distributions, and between sites (Region) for Long Island and Georges Bank of the same sex and between population age-frequency distributions. Tests include the Anderson–Darling (AD), Wald–Wolfowitz Runs (Runs), and Kolmogorov–Smirnov (KS) tests with a 0.05 significance level (non-significant results [ns]). Georges Bank values were compiled from Hemeon et al. (2021a). Region compared sex-specific (male [M] and female [F]) distributions between Georges Bank and Long Island

the shortest life expectancy (Table 4; Fig. 9). Mortality rate estimates produced by the Ricker method were consistently higher at GB than at LI.

Mean longevity estimates derived from the simulated Ricker linear regressions were used in the Then et al. (2015) formula for t_{max} to provide an alternative estimate of mortality. This method resulted in more conservative estimates of mortality at GB, but higher mortality estimates at LI when compared to the mean mortality estimates derived from the Ricker method (Table 4). The trend of lower male mortality and higher female mortality, coupled with higher mortality at GB versus LI, persisted regardless of which mortality estimator method was chosen. As the fishery fully selects for the size classes used for these analyses, the influence of fishing is unlikely to have biased the LI estimates due to lower mortality estimates compared to GB, despite the importance of the LI region to total landings.

Sex ratios

A significant sex ratio pattern across 5-mm size classes was identified at LI: males dominated size classes 65 mm-84 mm and females dominated size classes 95 mm-114 mm (Table 7). Between 85 mm and 94 mm, dominance transitioned from males to females and the two sexes were at an approximately 1:1 (F:M) ratio. In other words, the sex ratios of the 85-mm and 95-mm size classes were not significantly different than a 0.5 expected ratio using the binomial
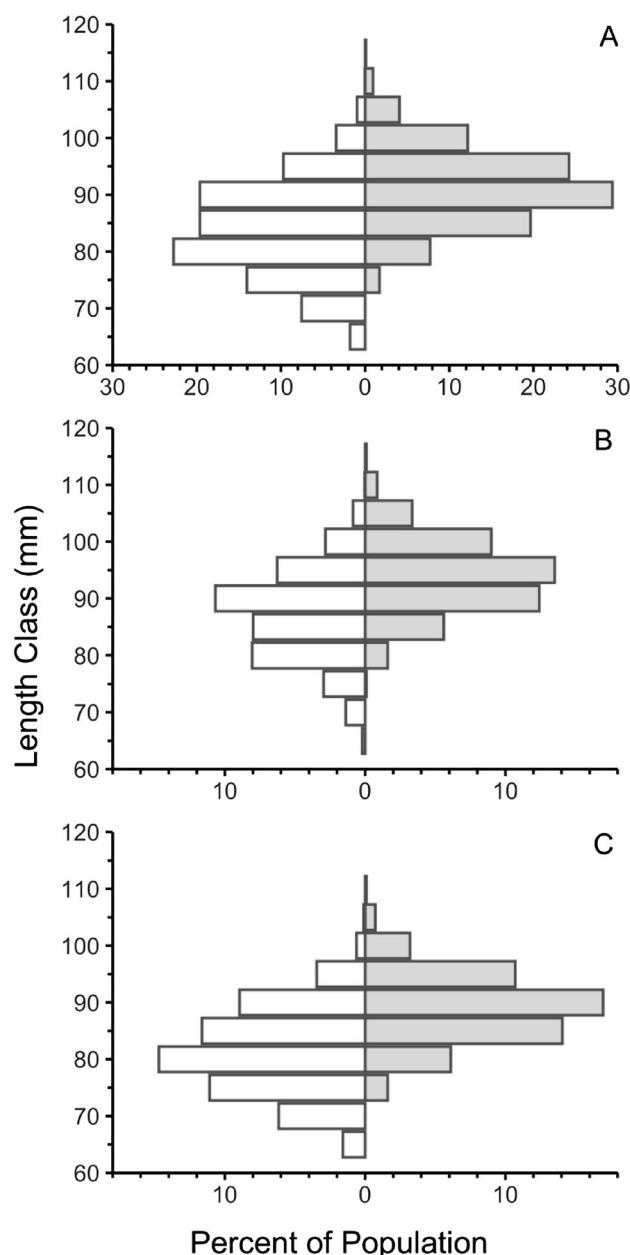
Fig. 4 Regional length-frequency data. Long Island (white) vs Georges Bank (grey) length frequencies as a percent of the total site-specific population by 5-mm size class of the **A** population, **B** female, and **C** male groups

test (Table 7; Fig. 10). The population sex ratio available to the fishery is 1:1.4 (F:M) and is significantly dominated by males despite similar mortality rates between sexes for fully selected age classes (Table 7). A two-sample approximate binomial test identified that sex ratios were significantly different between LI and GB at the 80-, 85-, 95-, and 105-mm size classes.

Table 6 Long Island summary table of all aged *Arctica islandica* sampled in 2017

Size class (mm)	N[a]	Female	Male	Mean age (y)	SD[b] (y)	Range (y)
60	3	1	2	54	32	63
65	30	4	26	43	15	49
70	78	15	63	48	21	141
75	97	32	65	63	17	96
80	125	63	62	77	23	163
85	126	63	63	98	33	169
90	120	59	61	123	43	201
95	125	62	63	151	43	179
100	96	52	44	169	47	209
105	86	80	6	181	46	249
110	18	17	1	188	44	141
Mean	82	41	41	109	33	151

[a]Sample size, [b]Standard deviation

Discussion

Reliability of age–length keys

This study aims to compare the population demographics of *Arctica islandica* from two distinct management areas of the US Mid-Atlantic fishery. The two management areas are contiguous but delineated by the Great South Channel. If stock assessments move forward with integrating age data into the assessment models, one critical element is the application of an ALK to estimate ages from a length sample, as the majority of survey and landings data are solely represented by lengths. Determining whether a single ALK is sufficient to represent the entire stock or if multiple ALKs are required is crucial. If the latter, ascertaining at what geographic scale age–length dynamics vary, and therefore necessitate different ALKs becomes an imperative.

This study found that population ALKs created for LI and GB were sufficient to produce site-specific population age frequencies. A population age frequency would be a time- and cost-effective alternative to sex-specific age frequencies, since additional laboratory equipment and slide preparation would not be necessary to distinguish samples by sex. However, ALKs are distinct by sex and site and a single ALK cannot be used interchangeably between LI and GB without generating increased uncertainty in the population age frequency. Also, organizing age–length data by sex provides extensive information on local population dynamics and illuminates an unusual life history not common in marine bivalves.

Dimorphism

Female length distributions are consistently offset to larger sizes in comparison to males within populations at both GB and LI. Mann–Whitney–Wilcoxon tests identified that the mean female lengths are significantly larger than the males at LI, and KS and Runs tests identified that the length distribution modes are significantly different and length distributions are offset between males and females. Age compositions are not significantly different between males and females at GB by mean age or age-frequency distribution (Hemeon et al. 2021a). A large difference in size but not age between sexes of a species is a strong indicator of sexual dimorphism, as demonstrated at GB by Hemeon et al. (2021a). LI age compositions were significantly different between sexes by both mean age and age-frequency distribution. If protandry was the explanation for the size offsets between males and females, wherein males would transition to reproductively distinct females at a predetermined size, one would expect males to dominate the youngest ages and smallest sizes and females to dominate the oldest ages and largest sizes. In that both very old and small males exist alongside both small and very old females at LI and GB, and male and female length distributions are significantly offset, our findings support those of Hemeon et al. (2021a) that *A. islandica* are sexually dimorphic and agree with Ropes et al. (1984a, b) and Fritz (1991) that this dimorphism arises from differential growth and not protandry.

Observation of sexual dimorphism is reinforced when sex ratios at size are considered. A significant difference in sex ratio exists for *A. islandica* at GB, and for animals less than 85 mm and greater than 95 mm at LI. Males at both sites dominated small size classes, and females dominated large size classes. A knife-edge transition in dominance occurred at 95 mm for GB, whereas LI demonstrated a gradual
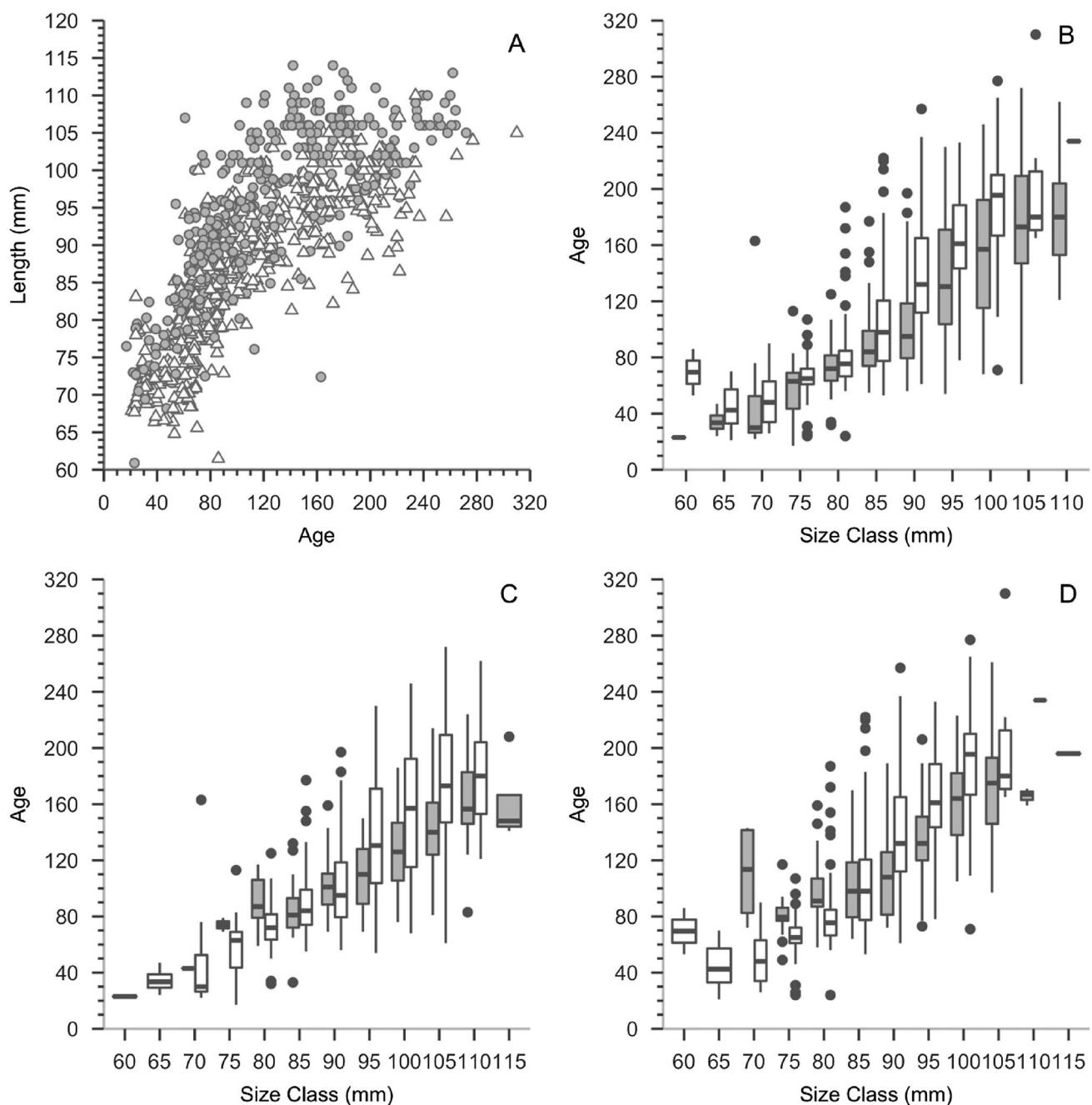
Fig. 5 Long Island and regional age–length data. Long Island aged data and comparisons with Georges Bank aged data. **A** Age at length data for Long Island female (grey) and male (white); **B** Long Island age compositions by 5-mm size classes between females (grey) and males (white); **C** female age compositions between Long Island (white) and Georges Bank (grey); **D** male age compositions between Long Island (white) and Georges Bank (grey). For boxplots, box represents the interquartile range (IQR) with 50th percentile bar (median), whiskers represent 1.5*IQR, and points are outliers

transition in dominance between 85 and 95 mm. Ropes et al. (1984a, b) collected samples in 1978 and 1980 at the onset of an *A. islandica* fishery in LI, New York (43,400 lbs. harvested between 1979 and 1980, NOAA 2021), and recorded a gradual transition of male to female dominance between 70 and 90 mm. A study by Thorarinsdóttir and Steingrímsson (2000) in northwest Iceland, where growth rates and maximum size are lower than those reported from the US continental shelf, identified the transition from male to female sex-ratio dominance at a smaller size of 40 mm. Maximum length and modal length distributions at LI are smaller than those at GB; it is unclear if the size discrepancy

Fig. 6 Long Island age frequencies by birth year. Standardized age frequencies as a percent of the total age frequency for the **A** population, **B** female, and **C** male datasets by birth year. Decrease in age frequency for recent years (~ 1960–2017) represents animals not fully selected by survey gear (gear highly selective for animals > 80 mm)
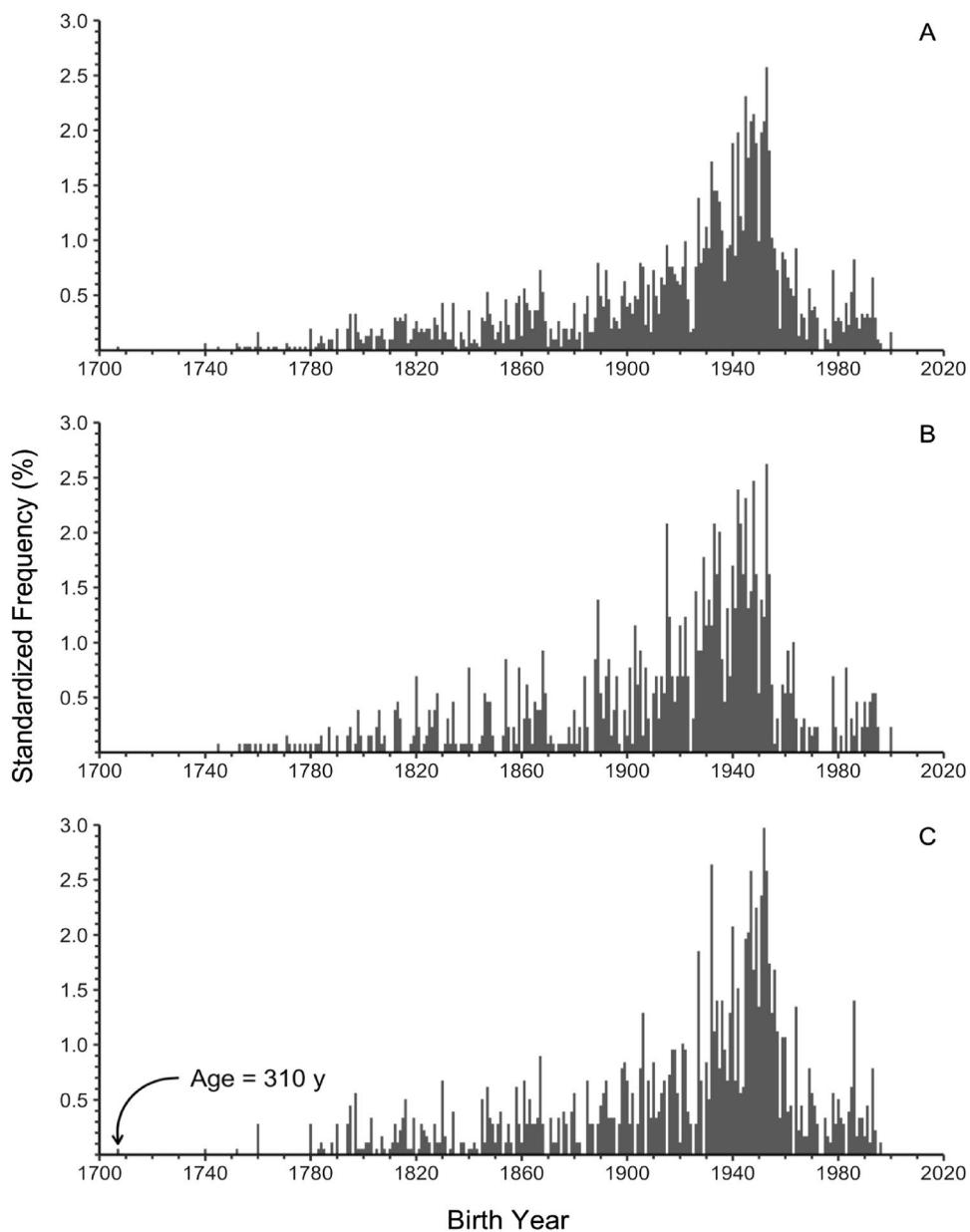
between sites is the result of reduced growth rates at LI or a size bias due to an active fishery, but a sex-ratio transition at a smaller size than observed at GB could indicate lower growth rates at LI as observed at northern latitudes (e.g., northwest Iceland). In other words, female *A. islandica* growth rates eclipse male growth rates at smaller sizes in colder/slower growing environments.

Fishery effects

The two *A. islandica* populations analyzed in this study were chosen to compare a relatively virgin population with a population supplying the greatest fishery landings in the Mid-Atlantic. Substantial observed differences between the

two populations may accrue as a consequence of the fishery and/or local oceanographic conditions. Long Island had a greater maximum observed age but smaller maximum and median lengths than observed at GB. The fishery at LI may bias the length distribution to smaller sizes but interestingly, no age bias (age truncation) attributable to the fishery appears to be present. Mortality rate estimates are lower at LI, a result unexpected if age truncation was present. However, the stock may be resistant to age truncation. Pace et al. (2018) observed increased growth rates over time in the Mid-Atlantic, in that recent cohorts reached large sizes faster than earlier cohorts. Thus, a fishery that targets large animals will increasingly harvest younger animals, thus limiting age truncation as an outcome.
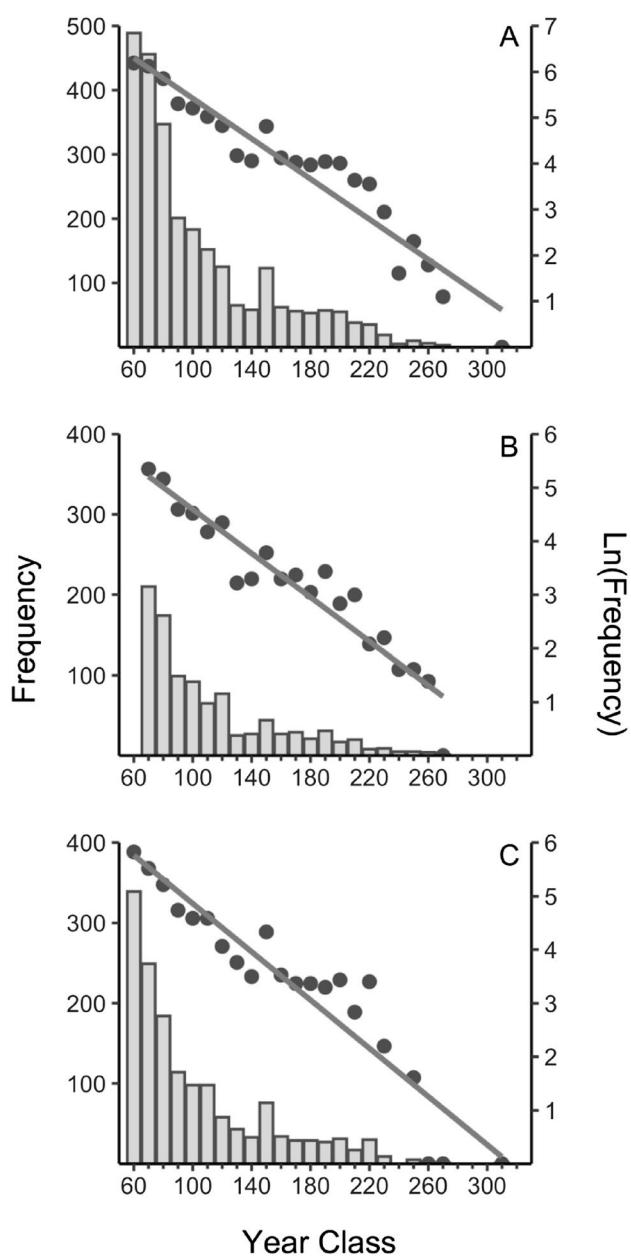
Fig. 7 Long Island mortality and longevity estimates. Total estimated mortality and longevity for Long Island **A** population, **B** female, and **C** male age-frequency data. Population longevity is 347 y, mortality is 0.02 ($R^2 = 0.91$); female longevity is 324 y, mortality is 0.02 ($R^2 = 0.90$); male longevity is 316 y, mortality is 0.02 ($R^2 = 0.89$)

Fig. 8 Regional mortality rates. Mean simulated mortality rates for Georges Bank (circle) and Long Island (triangle) with standard deviation, derived from female, male, and population age-frequency simulations
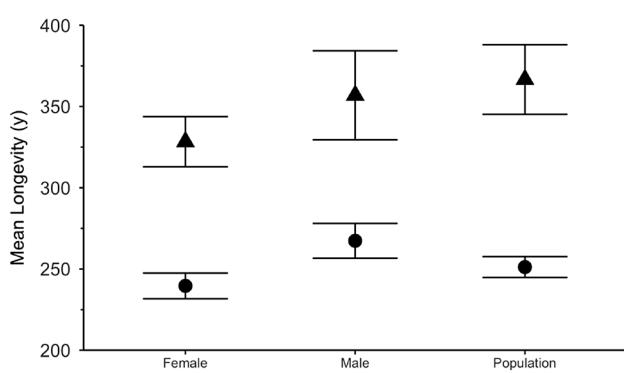
Fig. 9 Regional longevity estimates. Mean simulated longevity estimates for Georges Bank (circle) and Long Island (triangle) with standard deviation, derived from female, male, and population age-frequency simulations

A second distinction between GB and LI is the population sex ratios for animals > 70 mm. Hemeon et al. (2021a) discovered a 1:1.1 (F: M) sex ratio of the "unfished" GB population where male *A. islandica* were more available than females. A 1:1.4 population sex ratio was observed at LI, a fished population with an even more dramatic bias toward male *A. islandica*. Studies by Thorarinsdóttir and Steingrímsson (2000) in northwest Iceland and Rowell et al. (1990) in southwest Nova Scotia of unfished *A. islandica*

populations determined that the sex ratios were male dominated and aligned with results from Hemeon et al. (2021a); even fished populations tended to be dominated by males (Jones 1981; Mann 1982). In fact, Ropes et al. (1984a, b) is one of very few studies to find a population sex ratio biased toward females, but this study occurred at the onset of the LI *A. islandica* fishery. Conceivably, the decades-old fishery at LI may have fished down the largest *A. islandica*, presumably females, resulting in smaller maximum observed size and a length frequency shift to smaller size classes. The length compositions reported in NEFSC (2020) do not support this expectation, but these data do not overlap the Ropes (1984) timeframe; thus, a fishery that potentially targeted the large females of the population when the fishery first began could have resulted in the reduced overall size and female abundance at LI.

Table 7 Long Island sex ratios from the sex-determined shucked sample

Dataset	Sex ratio	Numbers			Female: Male
		N_T[a]	N_F[b]	N_M[c]	
Length sample	Population	2905	1205	1700	1: 1.40*
Shucked sample	60-mm	2	1	1	1: 1.00
	65-mm	26	4	22	1: 5.50*
	70-mm	84	16	68	1: 4.25*
	75-mm	116	27	89	1: 3.30*
	80-mm	193	69	124	1: 1.80*
	85-mm	219	96	123	1: 1.28
	90-mm	318	176	142	1: 0.81
	95-mm	397	267	130	1: 0.49*
	100-mm	303	249	54	1: 0.22*
	105-mm	107	98	9	1: 0.09*
	110-mm	21	20	1	1: 0.05*
	115-mm	2	2	0	–

Population sex ratio was determined from the length sample after selectivity adjustment. Sex ratios were evaluated using exact and approximate binomial tests, a significant result (*$P < 0.05$) indicated that the sex ratio was not 1:1

[a]Total number of samples, [b]Number of female samples, [c]Number of male samples
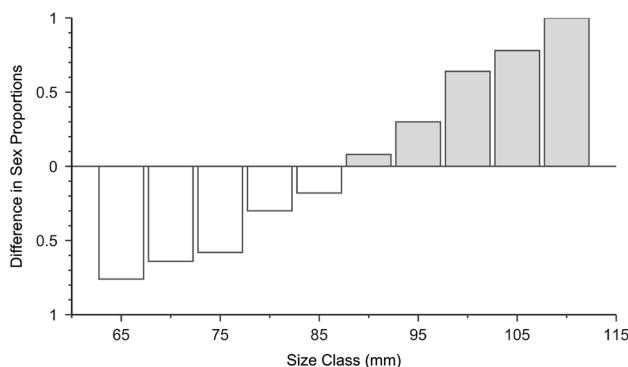
Fig. 10 Long Island sex ratios by size. Proportions by 5-mm size class of the shucked sample. The *x* axis (*y*=0) indicates a 1:1 sex ratio, white bars indicate a male-dominated class, and grey bars indicate a female-dominated class. The *y* axis denotes the proportional difference of the sex ratio, where 0.5 indicates a 1.5 to 1 sex ratio

The Cold Pool

The Cold Pool is an extremely important feature along the Mid-Atlantic Bight (MAB) that provides cold, low-salinity water to the MAB benthos as the spring/summer thermocline prevents benthic water from mixing with the warmer surface waters (Bigelow 1933; Brown et al. 2012). Cold water advances from the Gulf of Maine, around GB, and extends southward toward Cape Hatteras where the cold water remains within the MAB throughout the summer (Xu et al. 2015). The Cold Pool persists through the summer until advection warms the northern edges of the feature and September storms mix the warm surface water with the cold bottom water, whereby the northern boundary of the Cold Pool degrades faster than the southern boundary (Chen et al. 2018).

The fact that *A. islandica* shell accretion rates and growth trends are the result of both temperature and food supply is well established (e.g., Weidman et al. 1994; Schöne et al. 2005; Mette et al. 2016; Reynolds et al. 2016). Consequently, *A. islandica* populations adjacent to the northern boundary of the Cold Pool (e.g., LI) are subject to increased environmental variability as the Cold Pool footprint advances in the spring and retreats in the fall (e.g., fluctuating bottom water temperatures, surface water temperatures, food supply) versus GB that is relatively stable in comparison. Cold Pool seasonality is manifested as variability in *A. islandica* shell growth patterns that give rise to higher aging errors of LI samples than GB samples (see Sect. 3.1, Hemeon et al. 2021a,b). Using carbon-14 ages to calibrate aging methodology at LI (see Hemeon et al. 2021b), it was determined that LI *A. islandica* shells present substantially more subannual growth lines than those collected at GB. Subannual growth lines (sometimes referred to as check lines) are periods during a growth year when growth slows drastically in response to reduced metabolism and/or energy reallocation, and can be the result of a spawning event, sub-optimal temperatures, or inadequate food supply. Suppressed energy allotment for growth in the form of calcium carbonate formation results in the establishment of an additional growth line or subannual growth line (see Jones 1980; Chute et al. 2012). When subannual growth lines are present, growth lines must be critically interpreted for subannual growth lines to be distinguished from true annuli; a process that contributes to greater age-reader error. The Cold Pool clearly has an influence on growth rates as it supports different thermal ranges and thermal persistence periods than experienced outside of the Cold Pool footprint such as at GB.

Seasonal Cold Pool temperature variability may prolong cold bottom water temperatures and potentially reduce total mortality when lethal warm temperatures are persistent in other areas of the continental shelf such as at GB (Kavanaugh et al. 2017). This species can burrow into the sediment periodically and drastically lower metabolic functions when conditions are unfavorable; a process thought to promote longevity for this species (e.g., Strahl et al. 2011; Ballesta-Artero et al. 2019). Both extensive anaerobiotic capability and sulfide tolerance (Oeschger and Storey 1993) support an evolutionary adaptation for protracted burial. At LI, mortality rates are both equal between sexes and lower overall when compared to those at GB. Consequently, high mortality rates at GB are translated to shorter life

expectancies than LI where mortality rates are nearly half the rate of those observed at GB. It could be hypothesized that the warming temperatures at GB are lethal to some *A. islandica*, and burial may not be a sufficient escape for acute or chronic thermal exposure. The effect of the Cold Pool on total mortality may only be resolved with the addition of multiple Cold Pool sites with a comparison of mortality rates throughout the Southern Management Area, and a more thorough understanding of burrowing response to adverse environmental changes (see Ragnarsson and Thorarinsdóttir 2020).

Regional recruitment trends

The most apparent distinction between GB and LI is the lack of young animals in the observed size classes at GB and the extreme longevity at LI. Age–length data indicated that *A. islandica* at GB, regardless of sex, are younger at size than LI suggesting that GB growth rates are faster than growth rates at LI. If so, animals should reach the dredge selectivity minimum size at younger ages at GB than LI but contrary to expectation, very few young animals were aged at GB. Inconsistency between expected growth rates and observed ages may be the result of recent post-settlement mortality of animals within the past three decades, reduced recruitment over the past three decades, or patchy demographics at GB where smaller and potentially younger animals were not fully intermixed at the GB sample site. Benthic models developed for the US northeast continental shelf by Kavanaugh et al. (2017) estimated that between the years 1982 and 2015, benthic temperatures (depths less than 500 m) warmed faster than surface temperatures at GB across seasons, while the benthos off LI warmed slower than surface temperatures across seasons (see Figs. 6,7), with faster warming at GB than other offshore regions along the US continental shelf spanning Cape Hatteras to the Gulf of Maine. Recent benthic warming at GB compared to the Mid-Atlantic would support the hypothesis that recent benthic water temperatures at GB may be prohibitive to successful recruitment and increased post-settlement mortality of young *A. islandica* that are not able to burrow and escape thermal extremes.

At the older spectrum of the age-frequency distribution, extremely old animals were absent at GB compared to LI and could be the result of higher estimated natural mortality at GB. Mortality is often estimated from age frequencies assuming mortality is constant across time and age classes (Ricker 1975). Natural mortality estimates from GB age frequencies are higher than total mortality estimates at LI due to the extremely old ages discovered at LI that reduced mortality rates derived from a shallower regression model slope. If mortality is assumed to be stable over time and across ages, a higher mortality rate at GB could account for the lower representation of old animals.

The age frequency can also be viewed as a proxy for recruitment. Recruitment was consistent at LI between 1890 and 1970, in that animals were effectively recruited to the fishable size classes for each birth year during this time frame. Prior to 1890, animals were still born in most birth years but at lower frequency and with a few missing cohorts (Fig. 6). Proceeding from 1890, the age frequency began to steadily increase until its peak in 1955. The observed decline in the age frequency between 1955 and 2017 is likely a sampling artifact of younger animals not being fully available to the survey dredge. Harding et al. (2008) attributed bottom water temperature as a primary driver for divergent growth trends for this species, but local food availability is also considered an important factor. Mann et al. (2009) demonstrated a consistent warming trend that initiated in the early- to mid-1800s at the conclusion of the Little Ice Age (1400–1700 CE), a warming temperature trend could have driven the late-1800 increase in effective recruitment at both LI and GB (see Hemeon et al. 2021a).

In the 1970s, the LI population and sex-specific age frequencies declined dramatically and was followed by increased effective recruitment between 1980 and 1990 (see Fig. 6). Since many animals were born after the 1970s decline in the age frequency, the reduced number of animals surviving to the fishery from the 1970s is likely a true recruitment effect and not solely a result of low gear selectivity for these age classes. All age frequencies from LI also presented large decreases in effective recruitment during the mid-1920s. The Atlantic Multidecadal Oscillation (AMO) is an oceanographic cycle with recurring 60–80 y periodicity in the northern Atlantic Ocean and is predominantly characterized by positive (warm) and negative (cold) temperature indices (Alexander et al. 2014). The AMO had an observed negative index (negative temperature anomaly) in the mid-1970s that was comparable to the AMO negative index in the 1910–1920s (Nye et al. 2014), followed by a positive AMO index immediately succeeding in the 1990–2000s. The entire ~60 y period of an AMO cycle is not easily observed in an age frequency from either GB or LI, but extreme temperature anomalies do correspond with dramatic *A. islandica* effective recruitment events (e.g., lows 1920s, 1970s). Hemeon et al. (2021a) also observed substantial declines in the GB age frequencies during the 1920s, but very few young animals were sampled from GB to illuminate the presence of a 1970s recruitment decline during the sample time series. If the 1970s were a truly poor decade for *A. islandica* recruitment in the Mid-Atlantic, it may explain why very few young animals were observed at GB or even suggest a stronger 1970s negative-recruitment effect at GB compared to LI. Temporal coherence between AMO indices and effective recruitment trends of *A. islandica* support findings that extreme negative AMO indices (extreme cold temperature anomalies) produce years of reduced population recruitment at GB and LI. It is well known that climatic

cycles with strong bottom-water temperature variability affect *A. islandica* growth rates (Harding et al. 2008; Poitevin et al. 2019), and it would not be surprising if the same positive and negative thermal growth trends apply to positive and negative recruitment trends.

Hemeon et al. (2021a) identified 8-y recruitment signals in the GB age frequencies and theorized that apparently stronger year classes could be the product of the North Atlantic Oscillation (NAO) cycles (Visbeck et al. 2001; Soniat et al. 2006, 2009). High recruitment years at LI occurred in approximately 1953, 1945, 1942, 1932, 1927, 1922, 1915, 1905, and 1889 (Fig. 6), with a mean peak recruitment cycle of also 8 y. Along with 8-y periods of high recruitment, 8-y periods of low to very-low recruitment also are present, resulting in decades with extremely high recruitment followed quickly by extremely low recruitment within only a 3–5 y time span (Fig. 6). The peak recruitment years at GB and LI are often only different by a year or two between the two sites, with LI lagging behind the GB pulses. Despite localized differences in benthic conditions on either side of the Great South Channel separating GB from the LI/southern New England continental shelf, recruitment timing appears to be consistent in this Mid-Atlantic region and could be the result of underlying oceanographic cycles such as the NAO. These two sites are also tightly linked by the Labrador Current that carries Arctic water from the Labrador Sea through the Great South Channel and around the southern flank of GB past the Nantucket Shoals to the Mid-Atlantic Cold Pool (Chen et al. 2018; Chen and Curchitser 2020). The movement of Arctic water to GB and the Mid-Atlantic strongly influences bottom-water temperatures, and the lead/lag in recruitment cycles of GB and LI are potentially a reflection of the lead/lag relationship between water movement throughout this region. Lagged recruitment pulses at LI from GB would be appropriate considering the water mass movement from the Labrador Sea arrives at GB first, before moving southward and contributing to the formation of the Cold Pool off LI (Xu et al. 2015). Understanding the stable periodicity and drivers of successful recruitment can substantiate stock projections when other parameters are difficult to predict.

Summary

The US *Arctica islandica* stock is divided into two distinct area-specific assessment regions. One assessment model evaluates the northern region of the stock at GB, while the other model assesses a much larger area spanning the continental shelf west of GB southeast to southern Virginia. Two sites were evaluated in this study, GB located in the northern management area and LI located in the center of the southern management area. The GB site is relatively unfished and represents a pseudo-virgin population, whereas LI represents the greatest stock landings in the US Mid-Atlantic fishery. Currently, the assessment model applied to the active fishery in the southern region incorporates length-based data, but with capacity to incorporate age-based data if chosen.

Similarities between LI and GB include sexually dimorphic growth, a population sex ratio that is biased toward males, and relatively coherent recruitment cycles that occur in ~8-y periods. However, that is where the likenesses end. The modes of the length frequencies were offset by sex and site, with GB length frequency central tendencies generally larger than those of the LI length frequencies. Growth rates were faster at GB than LI, and female growth rates were faster than males. Not surprisingly, the ALKs at GB and LI were not interchangeable due to the above-mentioned age and length relationships and at least two independent keys are needed for the Mid-Atlantic stock. Female growth rates are divergent from males, and aging only females at a site to reduce aging error would not be sufficient to replicate the population ALK and subsequent population age frequency, both male and female ages are required. Total mortality rates at GB were approximately double those at LI, and GB appears to have a higher female mortality rate than males; although, GB had a smaller sample size than LI and GB mortality may be inflated due to the absence of very rare, very old animals in the age sample. Finally, the transition from male to female dominance by size class occurred across different size ranges dependent on the population, and this metric may interfere with predicted spawning-stock biomass estimates if the sex proportions are not designated correctly across a length sample by site.

Evidence listed herein asserts that the GB and LI sites require different ALKs, mortality estimates, and sex ratios at size. It is also still unknown if the LI population dynamics are comparable to other *A. islandica* populations within the southern management area, or if ALKs should be developed latitudinally or in reference to location in the Cold Pool footprint. Furthermore, additional data are needed to understand how patchy the age–length demographics are within and between populations, and if replication of the modal section of the age-frequency distribution is sufficient for the purposes of the assessment model, or if the tails need to be resolved in the ALK analyses.

Interpretation of an age-frequency distribution is a confounding task as an age frequency can be evaluated through either a mortality or recruitment lens. For instance, are the peaks and troughs of an age distribution the result of fluctuating environmental mortality that periodically removed animals from a population, or the product of successful recruitment over time that systematically added animals to a population? Each option assumes that the other variable is held constant through time. In reality, both mortality and recruitment rates likely change over time and survival and mortality are intrinsically linked, particularly for benthic

invertebrates that are highly dependent on external conditions for suitable habitat and reproductive success. When considering the age frequency of *A. islandica*, is the long tail of old animals born prior to 1890 the result of (1) low mortality rates long ago that allowed such old animals to persevere for centuries, (2) low recruitment during 1700–1890 at the initiation of a range expansion into the Mid-Atlantic 200–300 y ago followed by a population explosion in the 1880s, or (3) the product of successful genotypes that have optimal functionality and thus low mortality rates within these environments? Understanding conditions that drive mortality and recruitment in recent generations, can facilitate our collective interpretation of Mid-Atlantic shelf ecology in prior centuries as well as the concurrent multi-centenarians still living in the modern fishery.

Declarations

References

Alexander MA, Halimeda Kilbourne K, Nye JA (2014) Climate variability during warm and cold phases of the Atlantic Multidecadal Oscillation (AMO) 1871–2008. J Mar Syst 133:14–26. https://doi.org/10.1016/j.jmarsys.2013.07.017

Ballesta-Artero I, Augustine S, Witbaard R, Carroll ML, Mette MJ, Wanamaker Alan D, van der Meer J (2019) Energetics of the extremely long-living bivalve *Arctica islandica* based on a Dynamic Energy Budget model. J Sea Res 143:173–182. https://doi.org/10.1016/j.seares.2018.09.016

Bigelow HB (1933) Studies of the waters on the continental shelf, Cape Cod to Chesapeake Bay. Pap Phys Oceanogr Meteorol 2:94. https://doi.org/10.1575/1912/1144

Brown W, Boicourt W, Flagg C, Gangopadhyay A, Schofield O, Glenn S, Kohut J (2012) Mapping the Mid-Atlantic Cold Pool evolution and variability with ocean gliders and numerical models. In: 2012 Oceans. IEEE https://doi.org/10.1109/oceans.2012.6404970

Chen Z, Curchitser E, Chant R, Kang D (2018) Seasonal variability of the Cold Pool over the Mid-Atlantic Bight continental shelf. J Geophys Res Oceans 123:8203–8226. https://doi.org/10.1029/2018jc014148

Chen Z, Curchitser EN (2020) Interannual variability of the Mid-Atlantic Bight Cold Pool. J Geophys Res Oceans 125: e2020JC016445. https://doi.org/10.1029/2020jc016445

Chute AS, Wainright SC, Hart DR (2012) Timing of shell ring formation and patterns of shell growth in the sea scallop *Placopecten magellanicus* based on stable oxygen isotopes. J Shellfish Res 31:649–662. https://doi.org/10.2983/035.031.0308

Conover WJ (1980) Practical nonparametric statistics. New York: John Wiley and Sons. 493 pp. https://doi.org/10.1002/bimj.19730150311

Engmann S, Cousineau D (2011) Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff [sic] test. J Appl Quant Methods 6:1–17

Fritz LW (1991) Seasonal condition change, morphometrics, growth and sex ratio of the ocean quahog, *Arctica islandica* (Linnaeus, 1767) off New Jersey, U.S.A. J Shellfish Res 10:79–88

Harding JM, King SE, Powell EN, Mann R (2008) Decadal trends in age structure and recruitment patterns of ocean quahogs *Arctica islandica* from the Mid-Atlantic Bight in relation to water temperatures. J Shellfish Res 27:667–690. https://doi.org/10.2983/0730-8000(2008)27[667:dtiasa]2.0.co;2

Hemeon KM, Powell EN, Pace SM, Redmond TE, Mann R (2021a) Population dynamics of *Arctica islandica* at Georges Bank (US): an analysis of sex-based demographics. J Mar Biol Assoc UK 101:1003–1018. https://doi.org/10.1017/s0025315422000030

Hemeon KH, Powell EN, Robillard E, Pace SM, Redmond TE, Mann RL (2021b) Attainability of accurate age frequencies for ocean quahogs (*Arctica islandica*) using large datasets: protocol, reader precision, and error assessment. J Shellfish Res 40:255–267. https://doi.org/10.2983/035.040.0206

Hennen DR, Mann R, Charriere N, Nordahl VA (2016) Testing the performance of a hydraulic clam dredge modified to capture small animals. National Oceanic and Atmospheric Administration Technical Memorandum NMFS-NE-237, p 15

Jones DS (1980) Annual cycle of shell growth increment formation in two continental shelf bivalves and its paleoecologic significance. Paleobiology 6:331–340. https://doi.org/10.1017/s0094837300006837

Jones DS (1981) Reproductive cycles of the Atlantic surf clam *Spisula solidissima*, and the ocean quahog *Arctica islandica* off New Jersey. J Shellfish Res 1:23–32

Kavanaugh MT, Rheuban JE, Luis KMA, Doney SC (2017) Thirty-three years of ocean benthic warming along the U.S. northeast

continental shelf and slope: patterns, drivers, and ecological consequences. J Geophys Res Oceans 122:9399–9414. https://doi.org/10.1002/2017jc012953

Mann R (1982) The seasonal cycle of gonadal development in *Arctica islandica* from the southern New England shelf. Fish B-NOAA 80:315–326

Mann ME, Zhang Z, Rutherford S, Bradley RS, Hughes MK, Shindell D, Ammann C, Faluvegi G, Ni F (2009) Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. Science 326:1256–1260. https://doi.org/10.1126/science.1177303

Mette MJ, Wanamaker AD, Carroll ML, Ambrose WG, Retelle MJ (2016) Linking large-scale climate variability with *Arctica islandica* shell growth and geochemistry in northern Norway. Limnol Oceanogr 61:748–764. https://doi.org/10.1002/lno.10252

Murawski S, Ropes JW, Serchuk FM (1982) Growth of the ocean quahog, *Arctica islandica*, in the Middle Atlantic Bight. Fish Bull 80:21–34

NEFSC (2020) Stock assessment of the ocean quahog for 2020. Woods Hole, MA: Northeast Fisheries Science Center. NEFSC Ref. Doc.22–09. 210 pp

NOAA (2021) Annual Commercial Landing Statistics. www.fisheries.noaa.gov/national/sustainable-fisheries/commercial-fisheries-landings (Accessed 18 Feb 2022)

Nye JA, Baker MR, Bell R, Kenny A, Kilbourne KH, Friedland KD, Martino E, Stachura MM, Van Houtan KS, Wood R (2014) Ecosystem effects of the Atlantic Multidecadal Oscillation. J Marine Syst 133:103–116. https://doi.org/10.1016/j.jmarsys.2013.02.006

Oeschger R, Storey KB (1993) Impact of anoxia and hydrogen sulphide on the metabolism of *Arctica islandica* L. (Bivalvia). J Exp Mar Biol Ecol 170:213–226. https://doi.org/10.1016/0022-0981(93)90153-f

Pace SM, Powell EN, Mann R, Long MC (2017a) Comparison of age-frequency distributions for ocean quahogs *Arctica islandica* on the western Atlantic US continental shelf. Mar Ecol Prog Ser 585:81–98. https://doi.org/10.3354/meps12384

Pace SM, Powell EN, Mann R, Long MC, Klinck JM (2017b) Development of an age-frequency distribution for ocean quahogs (*Arctica islandica*) on Georges Bank. J Shellfish Res 36:41–53. https://doi.org/10.2983/035.036.0106

Pace SM, Powell EN, Mann R (2018) Two-hundred year record of increasing growth rates for ocean quahogs (*Arctica islandica*) from the northwestern Atlantic Ocean. J Exp Mar Biol Ecol 503:8–22. https://doi.org/10.1016/j.jembe.2018.01.010

Pettitt AN (1976) A two-sample Anderson-Darling rank statistic. Biometrika 63:161–168. https://doi.org/10.1093/biomet/63.1.161

Poitevin P, Thébault J, Siebert V, Donnet S, Archambault P, Doré J, Chauvaud L, Lazure P (2019) Growth response of *Arctica islandica* to North Atlantic oceanographic conditions since 1850. Front Mar Sci 6:483. https://doi.org/10.3389/fmars.2019.00483

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1989) Numerical recipes. Cambridge: Cambridge University Press. 702 pp. https://doi.org/10.1017/s0263574700010675

Ragnarsson S, Thorarinsdóttir GG (2020) Burrowing behaviour in ocean quahog (*Arctica islandica*) *in situ* and in the laboratory. Mar Freshw Res Inst: ISSN 2298–9137

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reynolds DJ, Scourse JD, Halloran PR, Nederbragt AJ, Wanamaker AD, Butler PG, Richardson CA, Heinemeier J, Eiríksson J, Knudsen KL, Hall IR (2016) Annually resolved North Atlantic marine climate over the last millennium. Nat Commun 7:13502. https://doi.org/10.1038/ncomms13502

Ricker WE (1975) Computation and interpretation of biological statistics of fish populations. Bull Fish Res Bd Can 191:1–382

Ridgway ID, Richardson CA, Scourse JD, Butler PG, Reynolds DJ (2012) The population structure and biology of the ocean quahog, *Arctica islandica*, in Belfast Lough, Northern Ireland. J Mar Biol Assoc U K 92:539–546. https://doi.org/10.1017/s0025315411000154

Ropes JW, Jones DS, Murawski SA, Serchuk FM, Jearld A Jr (1984a) Documentation of annual growth lines in ocean quahogs, *Arctica islandica* Linne. Fish Bull 82:1–19

Ropes JW, Murawski SA, Serchuk FM (1984b) Size, age, sexual maturity and sex ratio in ocean quahogs, *Arctica islandica* Linné, off Long Island, New York. Fish Bull 82:253–267

Rowell TW, Chaisson DR, McLane JT (1990) Size and age of sexual maturity and annual gametogenic cycle in the ocean quahog, *Arctica islandica* (Linnaeus, 1767), from coastal waters in Nova Scotia, Canada. J Shellfish Res 99:195–203

Schöne BR, Fiebig J, Pfeiffer M, Gleβ R, Hickson J, Johnson ALA, Dreyer W, Oschmann W (2005) Climate records from a bivalved Methuselah (*Arctica islandica*, Mollusca; Iceland). Palaeogeogr Palaeoclimatol Palaeoecol 228:130–148. https://doi.org/10.1016/j.palaeo.2005.03.049

Soniat TM, Klinck JM, Powell EN, Hofmann EE (2006) Understanding the success and failure of oyster populations: climatic cycles and *Perkinsus marinus*. J Shellfish Res 25:83–93. https://doi.org/10.2983/0730-8000(2006)25[83:utsafo]2.0.co;2

Soniat TM, Hofmann EE, Klinck JM, Powell EN (2009) Differential modulation of eastern oyster (*Crassostrea virginica*) disease parasites by the El Niño-Southern Oscillation and the North Atlantic Oscillation. Int J Earth Sci 98:99–114. https://doi.org/10.1007/s00531-008-0364-6

Strahl J, Brey T, Philipp EER, Thorarinsdóttir G, Fischer N, Wessels W, Abele D (2011) Physiological responses to self-induced burrowing and metabolic rate depression in the ocean quahog *Arctica islandica*. J Exp Biol 214:4223–4233. https://doi.org/10.1242/jeb.055178

Then AY, Hoenig JM, Hall NG, Hewitt DA (2015) Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. J Mar Sci 72:82–92. https://doi.org/10.1093/icesjms/fsu136

Thorarinsdóttir G, Steingrímsson S (2000) Size and age at sexual maturity and sex ratio in the ocean quahog, *Arctica islandica* (Linnaeus, 1767), off Northwest Iceland. J Shellfish Res 19:943–947

Visbeck MH, Hurrell JW, Polvani L, Cullen HM (2001) The North Atlantic Oscillation: past, present, and future. Proc Natl Acad Sci USA 98:12876–12877. https://doi.org/10.1073/pnas.231391598

Weidman CR, Jones GA, Kyger (1994) The long-lived mollusc *Arctica islandica:* A new paleoceanographic tool for the reconstruction of bottom temperatures for the continental shelves of the northern North Atlantic Ocean. J Geophys Res 99:18305. https://doi.org/10.1029/94JC01882

Xu H, Kim H-M, Nye JA, Hameed S (2015) Impacts of the North Atlantic Oscillation on sea surface temperature on the Northeast US Continental Shelf. Cont Shelf Res 105:60–66. https://doi.org/10.1016/j.csr.2015.06.005