The cryptic population biology of *Chthamalus fragilis* Darwin, 1854 (Cirripedia, Thoracica) on the Atlantic coast of North America

John P. Wares1,2,*, Sam Crickenberger3,6,*, Annette F. Govindarajan4, James L. Hamrick5, Katelyn M. Skoczen2, Dorset W. Trapnell5,*, and David S. Wethey6,*,

1Odum School of Ecology, University of Georgia, Athens, Georgia, USA 30602;  
2Department of Genetics, University of Georgia, Athens, Georgia, USA 30602;  
3The Swire Institute of Marine Science and School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China;  
4Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA 02543;  
5Department of Plant Biology, University of Georgia, Athens, Georgia, USA 30602; and  
6Department of Biological Sciences, University of South Carolina, Columbia, South Carolina, USA 29208  
Correspondence: John P. Wares; email: jpwares@uga.edu  
(Received 3 June 2018; accepted 8 August 2018)

ABSTRACT

Better understanding of the mechanisms by which novel species arrive in marine communities is not only important for documenting such arrivals but can also lead to a better awareness of the abiotic and biotic limits on species distributions. Here we integrate spatial survey data, metapopulation hindcast modeling, and new genetic data from allozymes and mitochondrial loci to identify how cryptic diversity within the barnacle species *Chthamalus fragilis* Darwin, 1854 responds to the coastal environment to establish contemporary patterns of diversity and abundance. This is of interest because *C. fragilis* has long been recognized as novel in the southern New England intertidal community, having arrived in the late 1800s as the likely result of either human introduction or increased global temperature. Our results demonstrate an elevated and somewhat distinct genetic diversity for this species in southern New England relative to the rest of its distributional range. Our models explain this best as resulting from a “natural” expansion of the barnacle’s range late in the 19th century, with genetic connectivity between its northern and southern populations now hampered by a recruitment gap along the Delmarva Peninsula.

Key Words: barnacles, biogeography, hindcast models, phylogeography, population genetics

INTRODUCTION

Most extant biological communities have been affected in some way as exotic components of organismal or genomic diversity have expanded their spatial extent through natural or anthropogenic processes of colonization. To better understand the mechanisms of range expansion or introduction, many studies have sought to 1) identify the likely source(s) of new diversity, 2) explain how exotic diversity elements became distributed through their native ranges (Wares et al., 2005; Dawson et al., 2010; Govindarajan et al., 2017), 3) demonstrate how species ranges change as a result of recurrent introductions (Pringle et al., 2011; Tepolt & Palumbi, 2015), and 4) document the differential success of genotypes in the new range (Krueger-Hadfield et al., 2016, 2017).

An interesting class of novel diversity in an ecosystem consists of organisms with source ranges that are relatively close to the new geographic range; that is, the novel diversity elements may have originated within the same or an adjacent ecoregion (Spalding et al., 2007). For example, we can expect to find a number of instances of organisms that have shifted their ranges in the late 20th century, typically poleward or to higher elevation, as climate has changed (Dawson et al., 2010; Canning-Clode et al., 2011; Sunday et al., 2012). Such range shifts may be uncorrelated with distributional shifts in other taxa (Moritz et al., 2008; Dawson et al., 2010) and thus result in novel ecological interactions (Gilman et al., 2010). Proximity of source diversity does not preclude an active human role in moving that diversity.

Studies of species shifts, then, must include ways of distinguishing between the active anthropogenic transport of an organism,
e.g. as cargo or ballast, or an intentional introduction, to a new community, and the “natural” and often independent movement of organisms in response to climate change. Tracking the “leading edge” and “trailing edge” of species’ distributions is possible when range shifts are contemporary (Sunday et al., 2012), but it is difficult to distinguish between mere “presence” and full “establishment” of a species. The isopod Idotea metallica Bosc, 1802, for example, is known to be transported by the Gulf Stream on algal/seagrass rafts from the western Atlantic to the North Sea (Naylor, 1957), but only in recent decades have reproductive individuals become established in the colder waters of the North Sea (Franke et al., 1999; Gutow, 2003). The barnacle Tetraclita rubescens Nilsson-cantell, 1931, with larvae that have a tremendous potential for dispersal (Dawson et al., 2010), could in the past occasionally be found far outside of its established range, but only as habitats warmed and densities increased dramatically outside its historic range was this barnacle’s expansion deemed notable (and self-sustaining; Dawson et al., 2010). Even when range shifts or introductions are known to have happened in historic times, information on the dynamics of novel populations has been insufficient and indirect means are needed to evaluate what happened.

The barnacle Chthamalus fragilis Darwin, 1854 is distributed along the eastern coast of the United States from the north shore of Cape Cod, Massachusetts to the Texas coast in the Gulf of Mexico (Dando & Southward, 1980; Carlton, 2002). The northern range of C. fragilis overlaps with, and is in part defined by, the southern range of the barnacle Semibalanus balanoïdes Linnaeus, 1767 in and around Cape Cod and eastern Long Island Sound (Wethey, 2002). Both species have internal fertilization, brood their embryos and release feeding planktonic larvae. S. balanoides broods embryos from November to early spring, and larvae recruit from February to June depending on locality (Crickenberger & Wethey, 2018). Chthamalus fragilis produces several broods each spring and summer and larvae are planktonic from May to October (Lang & Ackenjusen-Johns, 1981; Wethey, 1984; Young, 1991). The interaction between these two species has been interesting to observe at the microhabitat scale inasmuch as S. balanoides harbors distinct genomic diversity in individuals that are frequently sheltered by algae or water versus those in more exposed habitats (Schmidt & Rand, 1999, 2001) where they are most likely to interact with (and outcompete) C. fragilis. It is quite likely that C. fragilis has been introduced to the southern New England region (Carlton 2002, 2011), as there are no records of it in this well-studied coastal region prior to 1898 (Fig. 1). Carlton (2002) allows for two possible mechanisms of arrival, either through climate-mediated range expansion to southern New England from the mid-Atlantic, or via fouled boat hulls or other transported materials from its native southern range. Wethey (1984) and Carlton (2002) proposed that the expansion or introduction of C. fragilis was associated with a reduced abundance of S. balanoides under warmer than usual temperature conditions, as S. balanoides primarily outcompetes C. fragilis by smothering them with recruits (Wethey, 1983, 1984). Govindarajan et al. (2015a) compared mitochondrial haplotypes of C. fragilis individuals from the Florida Keys to Cape Cod and showed that the genetic diversity exhibited by this species in southern New England is unlikely to have originated from farther south than Cape Hatteras, NC. The haplotypes fell into three well-supported phylogenetic clades, one of which was found only at northern sampling sites (Clade II; Govindarajan et al., 2015a). Although genealogical diversity within species is not unusual, recognition that some of that diversity has a restricted spatial range is considered a hallmark of selection playing a role (Sotka et al., 2004). Particularly for organisms with broad dispersal potential, the time over which a geographic pattern of genetic diversity breaks down should be very quick unless there are environmental factors influencing the fitness and persistence of distinct genetic backgrounds (Sotka et al., 2004; Pringle & Wares, 2007).

The mitochondrial signature observed by Govindarajan et al. (2015a) is consistent with some of the nuclear genomic signal as assessed by allozyme loci. Dando & Southward (1980) used a series of allozyme markers to distinguish C. fragilis from its sister taxon, C. proteus Dando & Southward, 1980, and noted in particular that two markers, mannose-6-phosphate isomerase (MPI) and glucose-6-phosphate isomerase [also known as phosphoglucose isomerase (PGI)], exhibited an allelic and genotypic shift that was notably larger than other loci surveyed at two sites in Massachusetts and North Carolina. MPI and PGI allozymes have been implicated repeatedly in microhabitat selection in marine environments (Schmidt et al., 2008). Most notably, the genotypes of MPI and PGI have been shown to separate individuals of S. balanoides between habitats that vary in low-tide exposure (Schmidt & Rand, 1999), and metabolic loci that exhibit high levels of heterozygosity have been shown to be likely targets of natural selection (Skibinski & Ward, 2004; Marden, 2013).

It is possible that the restricted geographic range of clade II and the occurrence of certain rare alleles or genotypes at these loci are indicative of the same population genomic process, for example, either reproductive isolation or genomic interactions. If nuclear allele frequencies and mitochondrial haplogroup frequencies are unlinked via fitness differences, gene interactions, or reproductive isolation, individual-level analyses of cytonuclear disequilibrium

---

**Figure 1.** Documented distribution of *Chthamalus fragilis* between 1851 and 1874 (Gould, 1841; Darwin, 1854; Leidy, 1855; Verrill & Smith, 1874) (A); between 1898 and 1916 (Sumner, 1909; Sumner et al., 1911; Pilbary, 1916) (B), and between 1979 and 2016 (Lang, 1979; Dando & Southward, 1980; Wethey, 1983; Jones et al., 2012; this study) (C). Black circles represent presence and white circles represent absence.
should indicate this. On the other hand, without knowing the haplotypes of the individuals assayed by Dando & Southward (1980), or the nuclear genotypes of individuals assayed by Govindarajan et al. (2015a), we cannot be sure of the dynamics that maintain this spatial pattern.

Our goal is to integrate new population survey data, ecological modeling, mitochondrial DNA sequence data, and allozyme data to refine our understanding of the historical ecology and recent evolutionary dynamics of C. fragilis across a broad latitudinal range. The survey data consist of abundance and recruitment data that identify a notable recruitment gap between the presumed native and exotic ranges of the species. We assess new samples of C. fragilis that span this habitat discontinuity as well as the canonical biogeographic boundary of Cape Hatteras (e.g., Hutchins, 1947; Hayden & Dolan, 1976) to fill gaps of mitochondrial diversity from Govindarajan et al. (2015a) and test for cytonuclear disequilibrium between the MPI and PGI alleles and the mitochondrial genotypes. Integrating these different lines of evidence allows a more comprehensive assessment of the likely history of range expansion of C. fragilis.

**METHODS**

**Distributional surveys of abundance and recruitment**

The distribution and abundance of Chthamalus fragilis was surveyed between 31 March and 6 April 2016 at 15 sites with hard/permanent substrata (i.e. sea walls, groynes, jetties, pier pilings) from Delaware to North Carolina, USA (Table 1, Fig. 1). Abundance was determined in 10 × 10 cm photographic quadrats using a Nikon Coolpix AW100 or Olympus TG-4 camera. Chthamalus fragilis was recorded as absent if it could not be found at a particular site within a 30 min survey period.

Recruitment was tracked in permanent quadrats at 5 sites within the extent of the distributional surveys (Jennette’s Pier, NC; Kitopeke State Park, VA; Cape Charles, VA; Indian River Inlet, DE; Lewes, DE). Eight 4 × 4 cm recruitment quadrats were scraped clean of adult C. fragilis between 6 and 10 May 2016 at each site and photographed twice, during the periods of 23–27 July and 15–16 December 2016, respectively, using an Olympus TG-4 camera attached to a camera framer with a 4 × 4 cm base. Jennette’s Pier, NC was not sampled in December. Densities of C. fragilis adults and recruits were quantified using ImageJ (Schindelin et al., 2015) as the number of C. fragilis per unit area of each quadrat. Adult and recruitment densities were converted to a semi-quantitative scale (as in Crisp & Southward, 1958; Jones et al., 2012) for mapping the distribution and abundance of C. fragilis.

**Semibalanus balanoides recruitment modelling, historical sea-surface temperatures, and Chthamalus metapopulation modelling**

To examine historical changes in the recruitment success of S. balanoides along the Atlantic coastline of the U.S., we modeled and mapped the barnacle’s fertilization success for each year between 1815 and 2013 by applying the relationship between temperature and fertilization rate reported by Crickenberger & Wethey (2017, 2018) to average sea surface temperatures (SST) for November of each year, these in turn being based on SODAsi 3 ocean reanalysis SST data (pixels 0.5° × 0.5°, SODAsi 3; Giese et al., 2016). November SST was used because fertilization of S. balanoides occurs between late October and November in the western Atlantic (Barnes, 1956; Barnes & Barnes, 1976; Yuen & Hoch, 2010). We then modeled and mapped barnacle recruitment for each year during the same period by applying the relationship between log of recruitment and winter temperature during the brooding time derived from the 33-year time series of Abernat-Le Gac et al. (2016). This relationship was mapped onto December SST data for each year from the SODAsi 3 SST data (1815 to 2013, pixels 0.5° × 0.5°, SODAsi 3; Giese et al., 2016) because larvae of S. balanoides are released into the plankton as early as December in the western Atlantic, making December the month for brooding (Fish, 1925; Barnes, 1956; Crisp, 1964; Barnes & Barnes, 1976; J. Pineda, personal communication). We then multiplied predictions from the fertilization (proportion fertilized) and recruitment (number of recruits 100 cm−2) maps to predict the geographic distribution of recruits 100 cm−2 annually for each pixel between 1815 and 2013 (see Crickenberger & Wethey, 2017 for details). We have used this approach to model centennial scale changes in the biogeography of S. balanoides along the western Atlantic coast of the U.S. (see Crickenberger & Wethey, 2017 for details).

**Table 1.** Sites in this study generating new information regarding diversity and abundance in Chthamalus fragilis. Density and recruitment data are reported further in Figure 2; allozyme data are reported further in Table 3; mitochondrial data (with counts by phylogenetic clades I/II/III) are described further in Table 1.

<table>
<thead>
<tr>
<th>Site Name (abbreviation)</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Density/recruitment</th>
<th>Allozymes</th>
<th>mtDNA (I/II/III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woods Hole, MA (WH)</td>
<td>41.5260</td>
<td>−70.6731</td>
<td>present</td>
<td>+</td>
<td>28/93/37</td>
</tr>
<tr>
<td>Lewes, DE (LE)</td>
<td>38.7911</td>
<td>−75.1584</td>
<td>+</td>
<td>−</td>
<td>18/3/3</td>
</tr>
<tr>
<td>Indian River Inlet, DE (IR)</td>
<td>38.6076</td>
<td>−75.0608</td>
<td>+</td>
<td>−</td>
<td>6/0/0</td>
</tr>
<tr>
<td>Ocean City, MD (OC)</td>
<td>38.3243</td>
<td>−75.0851</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Chincoteague Island, VA (CI)</td>
<td>37.9010</td>
<td>−75.4075</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Cape Charles, VA (CC)</td>
<td>37.2668</td>
<td>−76.0263</td>
<td>+</td>
<td>−</td>
<td>26/15/2</td>
</tr>
<tr>
<td>Kitopeke State Park, VA (KP)</td>
<td>37.1673</td>
<td>−75.9887</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Fort Monroe, VA (FM)</td>
<td>37.0022</td>
<td>−76.3029</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>East Ocean Ave., VA (EO)</td>
<td>36.9503</td>
<td>−76.2420</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>East Beach, VA (EB)</td>
<td>36.9306</td>
<td>−76.1828</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Lynnhaven Pier, VA (LP)</td>
<td>36.9135</td>
<td>−76.0778</td>
<td>−</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Rudee Inlet, VA (RI)</td>
<td>36.8304</td>
<td>−75.9671</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Duck Pier, NC (DP)</td>
<td>36.1823</td>
<td>−75.7503</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Kitty Hawk Pier, NC (KH)</td>
<td>36.1014</td>
<td>−75.7109</td>
<td>+</td>
<td>−</td>
<td>10/9/4</td>
</tr>
<tr>
<td>Jennette’s Pier, NC (JP)</td>
<td>35.9101</td>
<td>−75.5954</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>Oregon Inlet, NC (OI)</td>
<td>35.7715</td>
<td>−75.5284</td>
<td>+</td>
<td>−</td>
<td>17/5/2</td>
</tr>
<tr>
<td>Duke University Marine Laboratory (DU)</td>
<td>34.7166</td>
<td>−76.6723</td>
<td>present</td>
<td>−</td>
<td>26/1/6</td>
</tr>
<tr>
<td>Tybee Island, GA (TY)</td>
<td>32.0416</td>
<td>−80.9433</td>
<td>present</td>
<td>+</td>
<td>20/0/5</td>
</tr>
</tbody>
</table>
Sampling for mitochondrial diversity and cytonuclear disequilibrium

Specimens of C. fragilis were collected to refine our understanding of mitochondrial diversity and to assess cytonuclear disequilibrium in comparison to the allozyme data described below. In particular, to increase spatial coverage of COI diversity relative to Govindarajan et al. (2015a), additional specimens of C. fragilis were collected from two sites in Delaware, Cape Charles, VA (where Clade II of Govindarajan et al. (2015a) is absent), and found that cirral activity ceases at 10 °C. Cirral activity is necessary for acquiring food, and feeding is not possible at temperatures below this threshold. To examine historical changes in the geographic location of this threshold (10 °C for cirral activity), we plotted the number of months in each year when monthly mean SST was greater than 10 °C along the east coast of the U.S. using SODAs 3 ocean reanalysis SST data (SODAs 3; Giese et al., 2016).

To model the effect of spatial changes in the 10 °C activity threshold on the biogeography of C. fragilis, we used an age-structured metapopulation model (Wethey et al., 2011, 2016) in conjunction with the most recent high-resolution reanalysis of ocean temperatures (SODAs 3; Giese et al., 2016). The SODAs 3 ocean reanalysis coupled an atmospheric reanalysis (Compo et al., 2011) derived from the International Surface Pressure Databank version 4 for the period 1815–1930 (Cram et al., 2015) to an ocean model (Carlton & Giese, 2008) that assimilated sparse International Comprehensive Ocean-Atmosphere Data Set (Woodruff et al., 2011) ship-of-opportunity data (Giese et al., 2016). The monthly SODAs surface temperature data (0.5° × 0.5°, 1815–2013) were bilinearly interpolated onto a 0.25° × 0.25° grid for use in metapopulation modelling. In 1815, populations of 10,000 in each of five adult-year classes were placed on all ocean pixels on the 0.25° × 0.25° grid. Each year class was assumed to suffer 50% mortality per year. If there were seven or more months above 10 °C from August to July (i.e. at least seven months of barnacle feeding activity), we assumed that barnacle reproduction was possible, so we allowed each 2- to 5-year-old individual in that ocean pixel in that year to produce 1,000 larvae, 1% of which dispersed to adjacent pixels. We allowed no reproduction if there were fewer than seven months above 10 °C from August to July in a pixel of a particular year. No other data were imported into the model (e.g., Wethey et al., 2011, 2016). Each population (pixel) was limited to a maximum density of 10,000 individuals in each age class. The model was spun up with five years of 1815 data and then the sequence of conditions from 1815 to 2013 was followed. This method makes two additional simplifying assumptions: that larval transport rates are similar each year (e.g., Crickenberger & Wethey, 2018) and that transport is geographically symmetrical (Wethey et al., 2011). Only monthly mean current velocities were available from SODAs 3, so direct simulation of larval transport was not practical.

Mitochondrial DNA sequencing

For the WH, CC, and TY specimens of C. fragilis, 2 μl of the allozyme protein slurry was removed and deposited into a labeled centrifuge tube. To this slurry, 8 μl of 10% w/v Chelex-100 (Bio-Rad, Hercules, CA, USA) was added and nucleic acids were stabilized as in Guo & Wares (2015). Individuals from other locations were also prepared using a Chelex stabilization protocol using whole soma dissected from the animal test. Stabilized nucleic acid samples were PCR-amplified using the mitochondrial cytochrome oxidase I primers from Folmer et al. (1994) using the same thermal cycler conditions as in Govindarajan et al. (2015a). Successful amplicons were cleaned for sequencing using an Evos I digestion with Antarctic phosphatase and submitted for Sanger sequencing at Macrogen.

Sequences were edited and aligned using CodonCode Aligner (v6.0.2, CodonCode Corporation; www.codoncode.com); sites with a PHRED score of less than 20 were considered ambiguous. A phylogeny of these sequences was inferred using MrBayes (v2.2.3) with the same model/parameter settings as in Govindarajan et al. (2015a). Sequence data were also analyzed using DNASp (Lefranc & Rozas, 2005) to estimate summary statistics by site and indices of site differentiation. To estimate site differentiation, we used $S_{am}$ of Hudson (2000) and a permutational test with 1,000 replicates; we also calculated $G_{ST}$ for context with other studies.

Cytonuclear disequilibrium (CND)

For individuals of C. fragilis from CC and WH, where clade II individuals were common, allozyme data were scored and matched to the individual mtCOI haplotype determined from Sanger
RESULTS

Distributional surveys of abundance and recruitment

Adult abundance and recruitment of *C. fragilis* were high throughout the region surveyed with the exception of sites north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, which separates Chesapeake Bay from the Atlantic Ocean (Fig. 2). Adult abundance averaged 6.6 individuals cm⁻² from Cape Charles southward; the 4 locations surveyed north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, out the region surveyed with the exception of sites north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, which separates Chesapeake Bay from the Atlantic Ocean (Fig. 2). Adult abundance averaged 6.6 individuals cm⁻² from Cape Charles southward; the 4 locations surveyed north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, which separates Chesapeake Bay from the Atlantic Ocean (Fig. 2). Adult abundance averaged 6.6 individuals cm⁻² from Cape Charles southward; the 4 locations surveyed north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, which separates Chesapeake Bay from the Atlantic Ocean (Fig. 2). Adult abundance averaged 6.6 individuals cm⁻² from Cape Charles southward; the 4 locations surveyed north of Cape Charles, VA along the outer coast of the Delmarva Peninsula, which separates Chesapeake Bay from the Atlantic Ocean (Fig. 2).

**Figure 2.** Adult density and recruitment of *Chthamalus fragilis*. Mean density (± SE, N = 6) of adult *C. fragilis* at sites listed from north to south (A). Mean recruitment (± SE, N = 6) of *C. fragilis* in July 2016 (dark gray) and December 2016 (light gray) at sites listed from north to south (B): LE, IR, CC, KP, and JP were sampled for recruitment, but JP was not sampled in December. Relative abundance and distribution of adult *C. fragilis* (Wethey, 1983; 1984; Jones et al., 2012; this study) (C). Relative abundance and distribution of *C. fragilis* recruitment (Wethey, 1983; this study) (D). See Table 1 for site abbreviations. ACFORN density scales after Crisp & Southward (1958): A, abundant, C, common, F, frequent, O, occasional, R, rare, N, none. Recruit densities are based on a similar scale where A = 10–100 individuals cm⁻², C = 1–10 individuals cm⁻², F = 0.1–1 individual cm⁻², O = .01–0.1 individuals cm⁻², R = 0–.01 individuals cm⁻², and N = 0 individuals cm⁻².

Mitochondrial diversity

We obtained 336 new mitochondrial COI sequence fragments of *C. fragilis* for analysis (NCBI accession numbers MG675651–675879; MG767174–767198, MH028963–029037, MH188628–188634). Although some low-quality end regions were trimmed (overall final length of alignment 569 nucleotides; 52 polymorphic sites as analyzed), all sequences were of sufficient quality to score diagnostic sequence motifs (see below) for the three clades identified in Govindarajan et al. (2015a). Using data from the present study as well as those from Govindarajan et al. (2015a), we confirmed the phylogenetic distinctness of the three original clades through an analysis in MrBayes (Ronquist & Huelsenbeck, 2003) using the same model parameters as in Govindarajan et al. (2015a). The same clades were recovered with posterior probabilities of 0.99 (clade I), 0.93 (clade II), and 0.99 (clade III). These clades are identifiable using single nucleotide polymorphism (SNP) positions relative to the conserved (within Crustacea) amino acid motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2). Each SNP is individually > 98% predictive of clade; 329-R (a motif IRAEL near the 5' end of the sequenced region (Table 2).

From these sequence data, the pattern identified previously (Govindarajan et al., 2015a) can be refined. Previously, the frequency of clade II individuals of *C. fragilis* was 0 at/near...
Savannah, GA, and the prevalence at Cape Charles, VA, was ~0.272 (both with small sample sizes). Here we were able to show with greater confidence that the prevalence of clade II at Tybee Island (near Savannah) is 0 (N = 23; N = 32 with the Govindarajan data added); at Beaufort, NC, 1/32 (0.031); and north of Cape Hatteras at Oregon Inlet and Kitty Hawk, NC, substantially higher (0.299). The frequency of Clade II at our site in Virginia is now estimated at 0.348, and our prevalence estimate for this clade at Woods Hole (N = 159; N = 77 from the upper portion of the barnacle’s vertical distribution and 92 from the lower portion) is 0.594, with no statistical difference between the upper and lower samples (Sm = 0.493; P = 0.473). A binomial regression of the prevalence of clade II by latitude (Fig. 4) shows a statistically significant association (P < 2x10^-16).

Given the strong spatial signature of clade II in C. fragilis, the increasing frequency of this phylotype alone leads to elevated population differentiation statistics (FST = 0.136 in Govindarajan et al., 2015a). To separate this signal from finer-scale assessments of gene flow, we evaluated gene flow only within clade I data, as data for this clade were recovered in sufficient numbers in all regions (sets of sampled locations that are separated by > 100km of unsampled individuals; Table 1). Analysis of site differentiation using only data from clade I indicated hierarchical structure between our WH data and the other three regions (Delaware, CC (VA), and North Carolina). Overall, Hudson’s S_m was 0.371 (four populations; P < 0.001); when WH is excluded, S_m = 0.338 (three populations; P = 0.35). Pairwise values of GST between WH and other regions ranged from 0.002 to 0.008 within clade I diversity alone and were negative for all other comparisons (-0.001 to -0.004). Haplotype diversity was similar in all regional populations (WH 0.973, Delaware 0.996, CC 0.966, North Carolina 0.982), whereas nucleotide diversity (π) was considerably higher in WH (0.0145) than the other three regions (0.0079–0.0088).

**Allozyme diversity**

Four alleles were identified for the MPI locus, and nine alleles for PGI [Supplementary material Table S1]. The frequency of the most common alleles for MPI were very similar between the two time periods (Dando & Southward, 1980; present study) but this was not so for PGI. The frequency of the most common PGI allele at Woods Hole in 1980 was 0.84, whereas those for the most common PGI allele (PGI-4; in each case the scored allele is indicated in order of mobility) from Woods Hole in 2016 were 0.713 (lower sample) and 0.747 (upper sample) [Supplementary material Table S1]. At CC (VA) and TY (GA) the frequencies of PGI-4 were 0.730 and 0.719, respectively. The highest rare PGI allele prevalence at WH was only 0.066 in Dando & Southward (1980) whereas the prevalence of the second most common allele (PGI-7) was somewhat higher in our samples, e.g. 0.080 and 0.121 at WH (upper and lower samples, respectively) and 0.146 at TY (GA). Overall, both alleles were polymorphic at all four sites surveyed; the effective number of alleles across both loci was higher in the WH samples (1.60 in the lower intertidal; 1.55 in the upper) than in Virginia (1.46) or Georgia (1.46); this result also held when samples were analyzed via rarefaction although the sample sizes were very different. Nei’s genetic identity among all sites was very high (> 0.99) for all pairwise comparisons; overall Gst from the allozyme data was 0.014, considerably lower than the

### Table 2

Single nucleotide polymorphisms (SNPs) from aligned sequence data of Chthamalus fragilis used to operationally define the three clades. Each position (top row) is counted from the start of an IRAEL amino acid motif near the 5’ end of the coding sequence; flanking nucleotides are shown for reference. All of these “definitive” SNPs (each individually is > 98% predictive) are at the third codon positions and are silent polymorphisms. Where there is variation at these sites within each clade it appears to be homoplasious, e.g. not parsimony informative. PHRED scores of 20 or higher are required at diagnostic sites for classification.

| Clade I | T | T | C | A | C |
| Clade II | R | T | T | G | C |
| Clade III | T | C | T | G | T |
structure exhibited by mitochondrial data when all phylogroups were included (0.136; Govindarajan et al., 2015a) but comparable to that seen when diversity from only a single phylogroup was analyzed (see above).

The small size of individuals scored from TY left little tissue for mitochondrial gene sequencing so the allozyme results and sequence results from TY were from different individuals and were not included in the analysis of cytonuclear disequilibrium (CND). Despite the large number of individuals with both mtDNA and allozyme data (N = 201) that were separable into clade II (N = 108) versus clades I or III (N = 83), all statistics for CND between the mitochondrial lineage and particular alleles or particular genotypes at these two loci were not statistically significant (P >> 0.05, Supplementary material Table S1).

**DISCUSSION**

The results of our study provide new insights into the appearance of *Chthamalus fragilis* in southern New England in the late 19th century. The historical documentation of the distribution of *C. fragilis* prior to this time strongly indicates that the abundance of the barnacle changed dramatically in New England, whether by human transport or a change in environmental and/or biotic conditions (Fig. 1; Carlton, 2002; Wethey, 2002; Carlton et al., 2011). There are nevertheless intriguing components to our data that bear further consideration. Information from diverse analyses suggests the presence of an abundance/recruitment gap that is concordant with a zone of apparently limited regional gene flow and a mitochondrial cline whereas our modeling results appear to favor a natural expansion of *C. fragilis* into its current distributional range.

First, there seems to be a range discontinuity in the mid-Atlantic region. Although we do not have observational or genetic data from New Jersey, Pilsbry (1916) listed a specimen of *C. fragilis* that had been collected in Ocean City, NJ: “this species was found by Mr. Fowler growing on *Balanus* (currently *Semibalanus* balanoides), the largest specimens 5mm. in greatest diameter.” This note is of interest because it highlights the close ecological interaction between *C. fragilis* and *S. balanoides*, and also because the material reported by Pilsbry (1916) was collected very close to regions we surveyed (IR and LE in Delaware), where *C. fragilis* was nearly absent and showed effectively no recruitment in recent
years (Fig. 2). This species has been listed as “common” in surveys of coastal armoring (jetties and groynes) on the shores of both New Jersey and Long Island, NY (USACE, 2013; Edinger et al., 2014). Abundance records in this region are otherwise limited, but Wethey (1983, 1984, 2002) reported 100% cover of C. fragilis on examined substrates at Horse Island, CT, and adult densities of 1 cm⁻² at the northern range limit in Massachusetts, which are lower than, but similar to, the densities we found throughout North Carolina and Chesapeake Bay (Fig. 2). Recruitment density in Horse Island (Wethey, 1983) was also similar in magnitude to that found throughout these latter regions (Fig. 2). While further surveys of recruitment and abundance along sections of intermediate coastline are needed to more clearly define this dip in recruitment and abundance, our data suggest that some persistent barrier(s) is (are) negatively affecting the success of C. fragilis in the Delmarva region, except for the peninsula’s southern tip at Cape Charles, and could maintain low gene flow between sites north and south of this region. Our data imply low gene flow between the population of C. fragilis at Woods Hole and those from Delaware, Virginia, and North Carolina, with elevated G_ST and significant values of Hudson’s FST between WH and all other sites. This is particularly notable in that the signal is consistent across our allozyme data as well as a subset of our mitochondrial data, chosen to exclude the apparent environmental association of mitochondrial clade II (Fig. 4).

Studies of genetic diversity have indicated the potential for a connectivity gap between southern New England and more southerly parts of the U.S. Atlantic coast in other coastal taxa as well as the present species of barnacle (Wares, 2002; Sotka et al., 2003; Diaz-Ferguson et al., 2009; Altman et al., 2013). The intervening region of the Delmarva Peninsula and the New Jersey/New York coasts clearly merits closer attention with respect to the density, abundance, and recruitment of its marine diversity and the respective regulatory mechanisms. For example, larval transport through this region may be hindered. Along the southeastern and Middle Atlantic U.S. coasts larval transport is primarily wind- or buoyancy-driven and varies seasonally (Epifanio & Garvine, 2001; Blanton et al., 2003). Recruitment in C. fragilis appears to begin at least as early as July in the Middle Atlantic area (Fig. 2). It has been thought to begin later in New England (Wethey, 1983, 1984), but Govindarajan et al. (2015b) identified Chthamalus larvae collected in the water column already in March by an autonomous underwater vehicle near our WH site. Among them, 3 COI haplotypes of C. fragilis (NCBI KM649639-40, KM649660- KM649665) were recovered, including six that are characteristic of clade II. In March, buoyancy-driven flows dominate southward nearshore larval transport in the Middle Atlantic, with retention and onshore transport back to natal sites dependent on wind direction (Epifanio et al., 1989). Southward transport may contribute to the low recruitment success and abundance of C. fragilis south of Delaware Bay, where buoyancy-driven flows travel southward for over 100 km (Epifanio & Garvine, 2001).

Additional mitochondrial data collected here from Woods Hole (WH), Cape Charles (CC), and several previously unsequenced locations support the findings of Govindarajan et al. (2015a) and highlight another curiosity about the expansion and/or introduction of C. fragilis to southern New England, where it is typically highly abundant (Wethey, 1983, 1984). Given the distribution of the three distinct phylogroups of mitochondrial diversity, one in particular (clade II) is only found in moderate to high frequency from Cape Hatteras northward (Table 1, Fig. 4). This phylogroup is in fact found at much higher prevalences at sites in Rhode Island and Massachusetts than anywhere else in the historic range of C. fragilis, indicating either a founder event that dramatically shifted the frequencies of these phylogroups (with little recurrent gene flow to homogenize the regions) or an association of this phylogroup with environmental variables such as available substrate or temperature (Fig. 5).

Future work involving evaluation of microhabitat and experimental manipulation of the distinct phylogroups will be important to assess whether individuals harboring mitochondrial clade II (or interacting diversity from the nuclear genome) have environmental tolerances distinct from those of other C. fragilis. Southward (1965), as reported in O’Riordan et al. (2010), showed that C. fragilis exhibited little variation in high-temperature mortality across its range. With regard to low temperatures, however, Southward (1964) suggested that the poleward boundary for C. fragilis ought to be set by the water temperature below which the barnacles can no longer feed. Wethey (2002) and Crisp (1950), however, successfully transplanted C. fragilis and C. montagu northward beyond their current ranges of distribution in the USA and United Kingdom, respectively, and the transplants were able to survive and reproduce. These results suggest that human-assisted transport could account for the sudden appearance of C. fragilis in New England in 1909 (Carlton, 2002). Our hindcast analysis of coastal temperature and metapopulation modeling over the past two centuries, however, suggests that variation and overall increase in temperature and planktonic dispersal may explain arrival of C. fragilis in southern New England (Fig. 3B, C).

The establishment of C. fragilis in the northernmost part of its current range could also be associated with environmental shifts that have influenced the abundance and recruitment of Semibalanus balanoides (Fig. 5). The close relationship between success of S. balanoides and C. fragilis at a given location suggests that in years during which warming events limited the establishment of S. balanoides, C. fragilis could have been able to expand its population and perhaps even its range. Recent analysis of long-term data sets and population modeling have shown that in warmer years and in warmer microhabitats, the North Atlantic species of Chthamalus are less susceptible to competition with S. balanoides than is otherwise the case (Wethey, 1984, 2002; Poloczanska, et al., 2008). Our model of recruitment success in S. balanoides predicts that recruitment should have been low throughout the modern range of C. fragilis in the 1870s, and patchy with both low and high levels of recruitment throughout the region just north of the historical range (Fig. 3A). The coincidence of this lull in recruitment success with increased traffic on shipping routes from the southeast to the northeast coast of the USA likely contributed to the colonization of New England by C. fragilis. Our metapopulation model also indicates that due to variation in ocean temperature conditions (Fig. 3B), the northern geographic limit of C. fragilis probably fluctuated between 37°N and 40°N from 1815 to 1860 and that the species likely moved north to the Woods Hole region (41.5°N) between 1870 and 1890 (Fig. 3C). This model is based on the simple assumption that C. fragilis needs at least seven months of feeding activity in order to reproduce successfully; considering the extreme simplicity of the model, it is remarkable how well it matches the historical data (Fig. 3C).

In addition to the recruitment gap apparent in our data (Fig. 2), one may question how the availability of suitable recruitment habitat has changed over time. Chthamalus fragilis is unusual among chthamalids in recruiting to cord-grass (Spartina) as well as rocks, docks, and other consolidated features (O’Riordan et al., 2010). Certainly, shoreline “hardening” has increased in the past century, although this has been of minor influence with respect to open-coast marshes (Gittman et al., 2013). Chthamalus fragilis has also been documented on gastropods (Crisp, 1990) and sea turtles (Frick et al., 1990). Some studies have shown that C. fragilis can tolerate reasonably low salinity (15–20 psu) (Gordon, 1969); nevertheless, it is uncommon in other estuarine habitats surveyed by the authors (JPW, personal observation). The species does not appear to use Spartina or even small cobbles as typical substrates in southern Massachusetts (A.G., personal observation; M. Bertain, personal communication).

Additional genetic markers are needed to fully understand both the history of range expansion in C. fragilis and the likely
 association of some of this diversity with cooler environments (Fig 4; Bay et al., 2018). Following concordant but subtle patterns in allele and genotype frequencies in C. fragilis (Dando & Southward, 1980), we added contemporary data from two allozyme loci to see whether they provide independent evidence of limits to gene flow and/or some form of cytonuclear disequilibrium with the mitochondrial diversity. Our data show no pattern of cytonuclear disequilibrium, suggesting that the three mitochondrial lineages are likely interbreeding, otherwise we would expect large allele frequency shifts in the nuclear loci among the mitochondrial lineages. This analysis also suggests no direct association between these metabolic loci and mitochondrial lineage. Data for these loci from a large number of individuals at CC (VA) and WH (MA) also suggest low genetic structure among these regions (Gst = 0.014), which is mirrored by the divergence seen in mitochondrial data (within clade I, to remove the influence of the strong pattern in clade II). Allelic and genotypic richness are higher in our WH samples than in those from CC (VA) or TY (GA) (see Table 2 and Supplementary material Table S1).

Altogether, it is not unusual for genetic data to present similar or even elevated diversity patterns in colonized versus source sites (Kolbe et al., 2004; Wares et al., 2005). The cryptic population biology of C. fragilis is somewhat typical of marine invertebrates whose spatial variation has been studied (Sotka et al., 2003, 2004; Waters et al., 2007; Kelly & Palumbi, 2010; Haye et al., 2014; DeBiase et al., 2015), including other chthamalid barnacles (Zardus & Hadfield, 2005; Tsang et al., 2008; Shemesh et al., 2009; Wares et al., 2009; Meyers et al., 2013; Ewers-Saucedo et al., 2016). What is intriguing in this instance is the combination of an apparent environmental association with some of the observed diversity in the context of range expansion or introduction. While the distribution of a species is never wholly static, we would like to use environmentally associated diversity to reconstruct historic events and predict the trajectory of populations given the influence of further climate change (Bay et al. 2018). Here we are left with compelling reason to believe that the diversity currently found in C. fragilis in southern New England is likely derived from neighboring Middle Atlantic populations north of Cape Hatteras, but that a combination of limited gene flow along the New Jersey and New York coasts, along with probable adaptive differences in this diversity, has generated the observed pattern of genetic diversity, even if we cannot definitively rule out anthropogenic transport as a cause for the appearance of C. fragilis in southern New England. How this informs our understanding of the maintenance of biogeographic and phylogeographic patterns in this same region (Wares, 2002; Diaz-Ferguson et al., 2009) will require further examination of this system.

SUPPLEMENTARY MATERIAL

Supplementary material is available at Journal of Crustacean Biology online.

S1 Table, Chthamalus allozyme data and cytonuclear disequilibrium analysis results for allozymes with mitochondrial haplotypes.

ACKNOWLEDGEMENTS

Significant effort towards data collection for this article by David Johnson, Danielle Doucette, Cecile Deen, Kat Clancy, Kerin Bentley, Eleanor Kunz, Katie Bockrath, Christine Evers-Saucedo, and Eric Holm is greatly appreciated. Paul Dando and Eve Southward both contributed interpretations of historical data and additional references. This manuscript was significantly improved through the input of three anonymous reviewers as well as Associate Editor Mark Grygier. This work was supported with funding from the University of Georgia Research Foundation and the UGA Department of Genetics to JPW, and by grants from NASA (NNX11AP77G) and NSF (1129401) to DSW.

REFERENCES


Downloaded from https://academic.oup.com/jcb/article-abstract/doi/10.1093/jcbiol/ruy070/5092668 by University of Georgia Libraries user on 10 September 2018