12-2017

Phylogeography of an Estuarian Calanoid Copepod; Acartia tonsa in the Texas Gulf of Mexico

Nicole J. Figueroa
The University of Texas Rio Grande Valley

PHYLOGEOGRAPHY OF AN ESTUARIAN CALANOID COPEPOD; *ACARTIA TONSA* IN THE TEXAS GULF OF MEXICO

A Thesis

by

NICOLE J. FIGUEROA

Submitted to the Graduate College of
The University of Texas Rio Grande Valley
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2017

Major Subject: Biology

PHYLOGEOGRAPHY OF AN ESTUARIAN CALANOID COPEPOD; *ACARTIA TONSA* IN THE TEXAS GULF OF MEXICO

A Thesis
by
NICOLE J. FIGUEROA

COMMITTEE MEMBERS

Dr. David Hicks
Chair of Committee

Dr. Daniel Provenzano
Committee Member

Dr. Erin Easton
Committee Member

Dr. Zen Faulkes
Committee Member

December 2017

ABSTRACT

Figueroa, Nicole J., Phylogeography of an Estuarian Calanoid Copepod; *Acartia tonsa* in the Texas Gulf of Mexico. Master of Science (MS), December, 2017, 51 pp., 2 tables, 8 figures, references, 35 titles;

The calanoid copepod, *Acartia tonsa* is one of the most abundant and well-studied estuarian species. However, the idea that this cosmopolitan species has unrestricted dispersal and high gene flow has been challenged. In this study, an evolutionary picture of the phylogeography of *A. tonsa* was developed using the mitochondrial gene cytochrome oxidase one (mtCOI). Multiple new lineages were found in the Texas Gulf of Mexico that are basal to northeastern Atlantic lineages. Connectivity was also observed between Brazil, the Texas Gulf of Mexico, and the northeastern Atlantic coast. The revised phylogeny shows a clear pattern of speciation as the species made a northward expansion since the last glacial cycle during the Pleistocene epoch. These data show that *A. tonsa* is a model species for observing phylogeographical structuring along the American continent.

DEDICATION

The completion of my Master's thesis would not have been possible without the unconditional support from my husband, Diego Figueroa who inspired, motivated and supported me in so many ways. I would not have accomplished this without you. Thank you for your love and patience.

ACKNOWLEDGMENTS

A special thanks to my thesis committee chair Dr. David Hicks who has given me support and advise throughout my career as a graduate student.

Thank you to my dissertation committee members: Dr. David Hicks, Dr. Daniel Provenzano, and Dr. Carlos Cintra whose comments and advice helped to increase the quality and accuracy of my work.

A special thank you to Dr. Diego Figueroa for teaching me molecular lab techniques and bioinformatic and for helping with the field work for this project. This project would not have been possible without the logistical support from Dr. Figueroa.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Significance of the Problem

Acartia tonsa (Dana, 1849) is a calanoid copepod (Copeopoda: Calanoida) in the family Acartiidae. It is a marine holoplankton commonly found in coastal estuaries. Marine holoplankton are organisms that have weak swimming capabilities and drift with water currents throughout their lives. *A. tonsa* is one of the most abundant and well-studied copepods in the world with a global distribution along the coasts of the Indo-Pacific and Atlantic (Blaxter et al., 1998). It is often the dominate copepod species in estuaries which makes it easy to collect. *A. tonsa* also plays a vital role in the trophic dynamics of most estuarian ecosystems as it is a main consumer of phytoplankton and a main food source for varied species of fish (Blaxter et al., 1998).

Because of life history characteristics and geographic distribution, *A. tonsa* was assumed to display low genetic diversity with a high potential for dispersal. Dispersal should not be hindered by geographical barriers due to the world's ocean circulation patterns resulting in worldwide distribution (Chen and Hare, 2008). Species such as *A. tonsa* are expected to have a large geographic range with few genetic differences or speciation between populations (Hirai et al., 2015). *Acarti tonsa,* and other estuarian specific species can experience geographical isolation because of the nature of their habitat. The geographical isolation due to residing in coastal estuaries has resulted in the divergence of multiple, distinct lineages of *A. tonsa* that in some cases live in sympatry (Chen and Hare, 2011, Caudill and Bucklin, 2004, and Costa et al, 2014). Four dis-

tinct lineages were discovered from the US Atlantic Ocean and the Gulf of Mexico (GOM) coasts using mitochondrial 16S rRNA (mt16S) sequences (Caudill & Bucklin, 2004). In Brazil (Costa et al., 2011), two independent lineages of *A. tonsa* were distinguished from mitochondrial gene cytochrome c oxidase I (mtCOI) sequences.

Chen and Hare (2008 and 2011) identified three lineages (F, X and S) of *A. tonsa* living in Chesapeake Bay and other Atlantic estuaries by using a partial sequence of mtCOI. These co-occurring lineages of *A. tonsa* were hypothesized to be separated by salinity gradients (Chen and Hare, 2008). *A. tonsa* persists in a wide range of salinities (0.3 to 36.9 psu) and temperatures (-1 to 35°C) (Lance, 1964; Gillespie, 1971; Gonzales, 1974).

Classifying and categorizing species can be challenging, especially when the definition of a species is purely philosophical. There is not a single species concept that is used consistently among biologists. Depending upon the taxonomic group that is being examined or the evolutionary processes that are being described, a different concept will be applied (Zachos, 2016 Pg 3). Some of the most frequently used species concepts are: The biological species concept states that a species is an inbreeding population that must produce viable, fertile offspring (Zachos, 2016 Pg 80). The evolutionary species concept defines a species as a lineage that has evolved separate from other lineages that also has its own evolutionary fate (Zachos, 2016 Pg 83). The genetic species concept refers to a population that is genetically isolated in terms of reproduction and whose individuals share genetic similarities (Zachos, 2016 Pg 87). There are many more species concepts (See Zachos, 2016) that can be applied to different organisms when classifying organisms. My study uses that evolutionary species concept to define a species.

Cryptic species, those that are morphologically similar but genetically different can pose challenges for taxonomists and biologists in terms of conservation, evolutionary theory, and biogeography (Avise et al., 1987; Bickford et al., 2007). Advances in molecular techniques and the ease with which we can identify species using molecular tools has greatly advanced the field of taxonomy. Because of the accessibility and availability of DNA sequencing, cryptic species have been identified in many taxa including algae (Saez et al., 2003), chaetognaths (Peijnenburg et al., 2004), copepods (Bucklin et al., 1996; Eunmi Lee, 2000; Caudill & Bucklin 2004; Goetze, 2003; Chen and Hare, 2008 and 2011; Costa, 2011), fish (Avise et al., 1987) and euphausiids (Zane et al., 2000). We can also learn a lot about the evolutionary history, patterns of distribution and impacts of geological and climatic events on an estuarine species by using molecular population genetic analyses (Avise et al., 1987; Caudill & Bucklin, 2004). For example, because of molecular data, Chen and Hare (2008) speculated that the evolutionary history of *A. tonsa* is associated with historical climate change events and geographic structure during the most recent ice age in the Pleistocene epoch (2.6 million to 11,700 years ago). The speciation pattern in the Texas GOM may be accounted for by the absence of any large-scale geological changes during the most recent ice age, when the Laurentide ice sheet extended as far south as 39° N (Jacobson et al., 1987, Dyke and Prest, 1987). The absence of such an event in the GOM lessens the chance for *A. tonsa* to have been separated by any reproductive barrier. By applying a crustacean mtCOI molecular clock, Chen and Hare (2011) hypothesized that the northern Atlantic lineages of *A. tonsa* diversified pre-Pleistocene and the mid-Atlantic lineage diversified post-Pleistocene. As ice expanded southward, holoplankton species would have been pushed southward or separated by falling sea levels and ice that caused geographical barriers. These barriers would have pro-

vided sufficient reproductive isolation of populations which had the potential to lead to speciation over time. As ice retreated and temperatures rose during warmer intervals of the Pleistocene, tropical species would have spread to higher latitudes (Chen and Hare, 2011). This evidence suggests that allopatric speciation only occurred in more northern latitudes. However, the processes and mechanisms underlying the diversification of lineages cannot be inferred without additional data from more southern sites, including the Texas GOM, which likely harbor distinct and potentially cryptic lineages from those identified by Chen and Hare, 2008 and 2011, Caudill and Bucklin, 2004, and Costa et al, 2014. Identifying lineages in the GOM may give a more complete picture of the evolutionary history of *A. tonsa*.

Rationale of the Project

Phylogeography examines the spatial distribution of gene lineages based on geographical, historical, and phylogenetic components (Avise, 2000). This phylogeography project uses sequences from GenBank and samples collected along the Texas GOM to compare haplotypes using mtCOI. These data were analyzed to show the phylogenetic relationship among *A. tonsa* lineages. Multiple genes including mtCOI and the nuclear internal transcribed spacer (nITS) have been used to establish lineages and clades in previous studies of *A. tonsa* (Caudill and Bucklin, 2004; Chen and Hare, 2008 and 2011; Costa et al., 2011). Chen and Hare (2008 and 2011) showed that there is congruence between mitochondrial and nuclear markers, so this study focuses on mtCOI.

In this study, I examined the genetic variation and phylogeographic structure of *A. tonsa* in the Texas GOM, Brazil, and the northeastern Atlantic coast. My expectations and hypotheses are as follows:

1. The phylogeographic structure of *A. tonsa* in the Texas GOM is expected to differ from that of the Atlantic. The Texas GOM has not been subjected to geological changes that could have caused

reproductive barriers that would have separated or killed off populations of *A. tonsa*, so basal lineages are likely to occupy the Texas GOM.

2. The S lineage described by Chen and Hare (2008) is expected to the dominant lineage in the Texas GOM because of its ability to tolerate higher salinity levels in Atlantic estuaries described by Chan and Hare (2008 and 2011).

3. I expect that any lineages belonging to Chen and Hare's (2008) F lineage found in the Texas GOM will be basal to the Atlantic lineages previously described by Chen and Hare (2011) as Caudill and Bucklin (2004) found a basal lineage to F5 in the Texas GOM. Further, I expect to find the same pattern of northward expansion within the F Clade as Chen and Hare (2011) described with more basal lineages in the GOM and a species expansion that follows the prevailing currents from the GOM to the northeastern Atlantic coast.

CHAPTER II

METHODOLOGY

Sample Collection

A. tonsa specimens were collected from July to November 2015 using a 300 µm simple plankton net either by hand tows from a fishing pier or by towing from a kayak. Three 100 meter tows were carried out at each of the six sites (Figure 1) to ensure at least ten *A. tonsa* specimens were collected. Specimens were immediately preserved in 95% ethanol in the field. Sorting and identification of *A. tonsa* individuals was accomplished using morphological analysis of the fifth leg and body shape characteristics with a dissecting and compound microscope by copepedologist Dr. Diego Figueroa.

Figure 1 Study sites in the Texas Gulf of Mexico. Site 1 - Bahia Grande (26° 1.103'N 97° 17.222'W), Site 2 - Port Mansfield (26° 34.176'N, 97° 25.660'W), Site 3 - Baffin Bay (27° 17.188'N, 97° 39.834'W), Site 4 - Port Aransas (27° 50.466'N, 97° 3.833'W), Site 5 - Port Lavaca (28° 38.321'N, 96° 36.722'W), Site 6 - Galveston Bay (29° 32.887'N, 95° 1.008'W).

Water Characteristics

The following water characteristics were measured at the site of specimen collection using a Eureka Manta 2 sub 3 multi-probe: temperature, pH, depth, phycoerythrin, chlorophyll a, dissolved oxygen, turbidity, and salinity (Table 1). It is important to note that environmental variables that were measured were only measured once at the time of sampling and they do not reflect a mean (particularly in terms of salinity) for the six sampling sites.

Table 1: Environmental variables and coordinates measured at the time of sampling for the six Texas Gulf of Mexico sites.

Site	Bahia Grande	Port Mansfield	Baffin Bay	Port Aransas	Port Lavaca	Galveston Bay
Latitude	26° 1.103'N	26° 34.176'N	27° 17.188'N	27° 50.466'N	28° 38.321'N	29° 32.887'N
Longitude	97° 17.222'W	97° 25.660'W	97° 39.834'W	97° 3.833'W	96° 36.722'W	95° 1.008'W
Temperature (°C)	31.37	26.71	26.14	25.53	24.83	23.5
pH	8.63	8.62	8.38	8.38	8.19	7.87
Depth (m)	0	1.39	0.82	1.52	0.53	0.13
Cyanobacteria (cells/mL)	262.2	428.7	301.3	151.7	509.2	290.6
Chlorophyll (ug/l)	17.23	26.73	10.81	8.29	36.37	8.37
Dissolved oxygen (mg/l)	8.75	7.08	6.5	6.42	6.94	7.78
Turbidity (NTU)	46.2	357.1	80.9	5.7	82.8	40.5
Salinity (psu)	32.4	30.8	29.4	30.2	13.6	2.6

Experimental Design

DNA Extraction and Molecular Analysis

In the laboratory, *A. tonsa* individuals were rehydrated in molecular-grade water for no less than 30 minutes. Genomic DNA was extracted by placing individuals in a 2 mL tube with 100 µl of Bio-Rad's Instagene Matrix. The specimens were then placed in a thermomixer and incubated at 56°C overnight. After incubation, samples were heated to 100°C for 8 minutes. The extracted DNA samples were stored in a -20°C freezer until further use. Quantification of extracted DNA was carried out using Thermofisher Scientific's Qubit fluorometer set to OD_{260}. Samples containing at least 0.1 ng/µl of DNA were subjected to polymerase chain reaction (PCR) to amplify the first half of the mtCOI gene known as the Folmer region which is a coding 658 base pair (bp) with primers LCO 1490 (5'GGTCAACAAATCATAAAGATATTGG3') and HCO 2198 (5'TAAACTTCAGGGTGACCAAAAAATCA3') (Folmer, 1994). Polymerase chain

reaction (PCR) was carried out in a 25 µl reaction: 7.55 µl PCR water, Invitrogen's 10X PCR Rxn Buffer (2.0 µl), 1.25 µl Invitrogen's 50 mM $MgCl_2$, 2.0 µl of 10 mM dNTP, 1.0 µl of 10 mM forward primer (HCO), 1.0 µl of 10 mM reverse primer (LCO), 0.2 µl Thermo Fisher's Invitro-gen Platinum TAQ DNA Polymerase, and 10.0 µl DNA. The following thermocycler conditions were employed: 94°C for 2 mins, followed by 40 cycles of 94°C for 1 min, 46°C for 1 min, 72°C for 1.5 mins, followed by 72°C for 1 additional min, and then cooled at 4°C ∞.

PCR products were visualized by agarose-gel electrophoresis followed by staining with 0.2 µg/mL ethidium bromide. A DNA ladder was electrophoresed with Life Technology's Invitrogen 1 KB Plus DNA ladder (1.0 µg/µL) with each gel to determine if a DNA band of the expected size (~700 base pairs (bp)) was amplified and to make sure other bands were not present. The PCR products with a single band of ~700 bp were purified using Sigma Aldrich's GenElute PCR clean-up kit. Purified PCR products were sequenced by Sanger sequencing technique by Eurofins MWG Operon LLC with the forward and reverse primers.

Bioinformatics

For each specimen, the sequences for the forward and reverse strands were aligned with the alignment program MUSCLE v3.8 (Edgar, 2004) with default parameters. MUSCLE was used as a plugin via CLC Workbench 7.9.1 (CLC Bio, Aarhus, Denmark). Chromatograms were visually inspected for conflicts between the two strands and conflicts were resolved manually. Base quality scores were visually examined for quality control and a consensus sequence was generated from the alignment. Sixty-three individual copepod COI Folmer regions of mtDNA were sequenced from the Texas GOM, and 319 sequences were downloaded from GenBank (National Center for Biotechnology Information (NCBI)[Internet] totaling 382 sequences (see Appendix). GenBank and consensus sequences were aligned with MUSCLE using default parameters (Edgar,

2004) and visually inspected for consistency. Each unique sequence in the MUSCLE alignment was treated as a haplotype. A list of 193 haplotypes was generated using the software dnaSP v5 (Librado, 2009). The haplotype list was generated as a NEXUS haplotype data file with gaps and missing sites considered and invariable sites included. The MUSCLE alignment was re-run with CLC Main workbench with default parameters using the 193 haplotype sequences. Phylogenetic analyses were performed using maximum-likelihood (ML) and Bayesian methods. Maximum Li-klihood analyses were performed with using PartitionFinder v1.1.1 (Lanfear, 2012) and RAxML v8.0.0 (Stamatakis, 2017) using the 193 haplotypes which were partitioned by codon. Blocks were defined by codon position and PartitionFinder found two partitions: The first partition was for codon positions one and two and the second partition was for the third codon position. For the Bayesian analyses, MrModeltest 2.3 (Nylander, 2004) determined the best fit model to be the HKY+I+G model of evolution. This model was used for the Bayesian analyses using Mr. Bayes 3.1 (Ronquist and Huelsenbeck, 2003). The two chains were carried out for 1,000,000 generations, sampling every 500th generation. After inspecting the trace files generated by the Bayesian Markov Chain Monte Carlo (MCMC) runs, the initial 25% (2,500) of sampled generations were omitted prior to building the consensus tree. Both phylogenies were rooted with *Acartia neglingens* (GenBank accession number EU856812) as an outgroup.

Haplotype Analysis

Haplotype network analysis was performed with PopArt v 1.7.2 with an epsilon of zero (Leigh & Bryant, 2015) using a median-joining network, which infers ancestral nodes by iteratively adding median sequence vectors (Leigh & Bryant, 2015). By using inferred ancestors, the PopArt software deduces relationships between haplotypes and provide a straightforward visual representation of those relationships. The MUSCLE alignments of the 193 haplotypes based on

the 557 bp region of mtCOI from CLC Workbench . If haplotypes with <557 bp are included in the median-joining network analyses, those missing base pairs will be ignored for all sequences in the analyses making it less accurate in determining relationships between haplotypes. For this reason, if a haplotype did not have at least 557 bp, it was removed from the network analyses, as was the case for the following six haplotypes that had shorter sequences: Haplotype one (423 bp), haplotype two (416 bp), haplotypes 26 and 28 (522 bp), haplotype 30 (451 bp) and haplotype 31 (473 bp). These haplotypes were not excluded from the ML or Bayesian analyses.

For the Texas GOM network analysis, the specific number of individuals per haplotype is known so the nodes are weighted by the frequency of individuals occurring in each haplotype (the larger the node, the more individuals belong to that specific haplotype). Because of the lack of such frequency data for the sequences used from GenBank, the network analyses that included these data were constructed based on a presence/absence method rather than weighting each node by the frequency of each haplotype.

Percent Divergence Between Haplotypes

Percent divergence shows the proportion of nucleotides that differ between haplogroups. Percent divergence between haplogroups was calculated by summing the branch lengths of the Bayesian tree from the phylogenetic analysis. The branch lengths illustrate the model-adjusted differences between each branch.

CHAPTER III

RESULTS

Phylogenetics

After trimming the MUSCLE alignment, resolving conflict between the two strands, and removing ambiguous base pairs from the ends of the sequences, the alignment of the haplotypes was 557 bp long. Sixty-three individual copepod mtCOI Folmer regions of mtDNA were sequenced from the Texas GOM, and 319 sequences were downloaded from GenBank totaling 382 sequences were assigned to 193 haplotypes, 38 of which are present in the Texas GOM. The phylogenetic reconstruction for *A. tonsa* shows six major lineages referred to as haplogroups I through VI with 6, 16, 39, 12, and 80 haplotypes, respectively. The ML and Bayesian analyses yielded a similar tree topology. The only differences between the two methods were that haplotypes eight and nine were placed in haplogroup IV using the Bayesian analysis and in haplogroup I using the ML tree (Figure 5) For this study, haplotypes eight and nine were placed in haplogroup IV because the Bayesian analysis had more robust results in terms of branch support when compared to the ML method. Bootstrap support and posterior probabilities from both analyses are indicated on the phylogenies (Figures 1, 2, 3, 4, 5, 6, and 7).

Percent nucleotide divergence between haplogroups ranges between 12.5-28.4%. Haplogroups IV and V display most homology with 12.5% divergence. Haplogroups I and IV are the most divergent haplogroups with 28.4% divergence (Table 3).

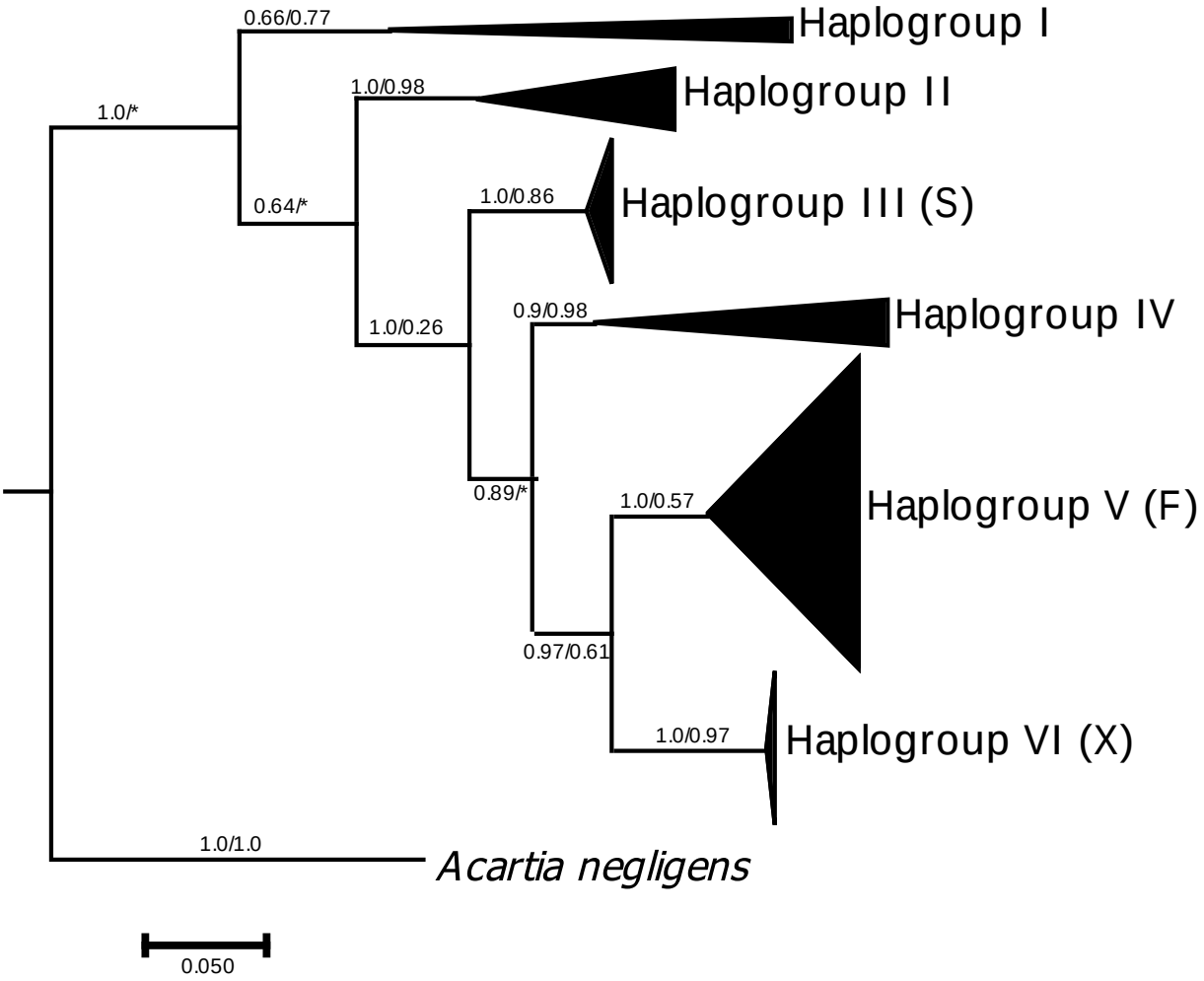

Figure 2 Phylogenetic tree based on Bayesian methods with collapsed branches showing the six major haplogroups of *Acartia tonsa* with their corresponding lineage names (F, X, and S) from Chen and Hare (2008 and 2011). Tree rooted with *A. negligens* sequence from GenBank (EU856812). Tree topology is supported by posterior probabilities from Bayesian analysis and bootstrap probabilities from ML values respectively. An * indicates bootstrap values < 0.50.

Haplogroup I

Haplogroup I is the basal lineage containing six haplotypes separated by geographic region: four from Rhode Island, one from California, and one from Port Aransas, Texas (Figure 2). The median-joining network analysis similarly revealed three distinct geographic groups within haplogroup I.

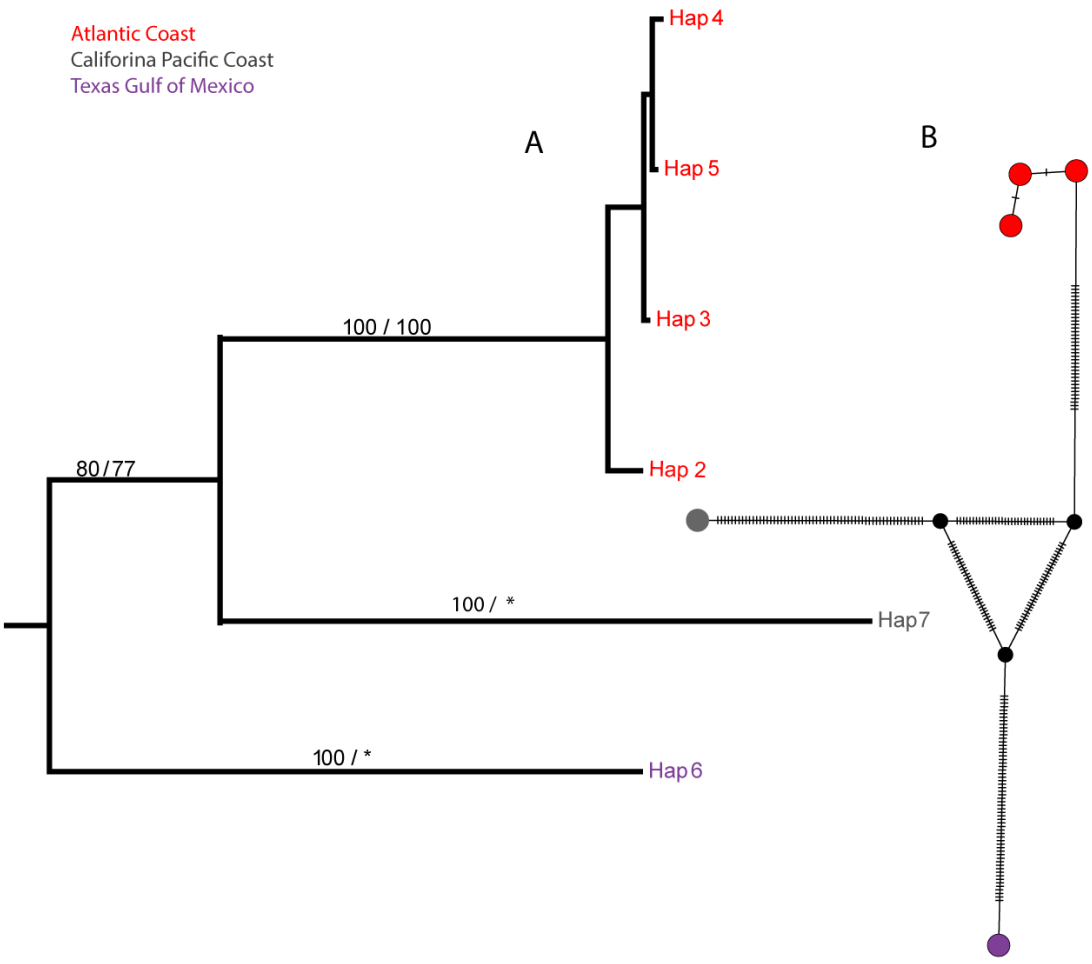

Figure 3 Phylogeographic analyses of Haplogroup I. The data was divided into three geographic groups: Texas GOM, California Pacific coast, and northeastern Atlantic coast. (**A)** Phylogeny of haplogroup I inferred by Bayesian methods. Branch values represent PP/bootstrap statistics. Tree topology based on ML methods was similar, an * was placed where bootstrap probabilities are < 0.50. Color of branches correspond to the geographic groups. (**B)** Median-joining network analysis of haplogroup I based on presence or absence of haplotypes. The size of the node does not reflect the haplotype frequency. The size of the nodes represents the number of regions where each haplotype is present. Haplotype 2 was excluded from the network analysis because it had <557 base pairs. The small, black nodes represent an inferred ancestral haplotype. Notches on each branch represent the number of nucleotide changes that occurred from one haplotype to another. The color of each node represents the location the individual was collected from (see legend).

Haplogroup II

The second lineage, referred to as haplogroup II, contains three clades. Two of the clades are from Brazil (Costa et al., 2011 and 2014) with three haplotypes and seven haplotypes; the third is a sister clade to these Brazil lineages with six haplotypes from the Texas GOM, all from Port Mansfield. The median joining network analysis similarly revealed three distinct groups one from the Texas GOM and two from the Brazilian coast.

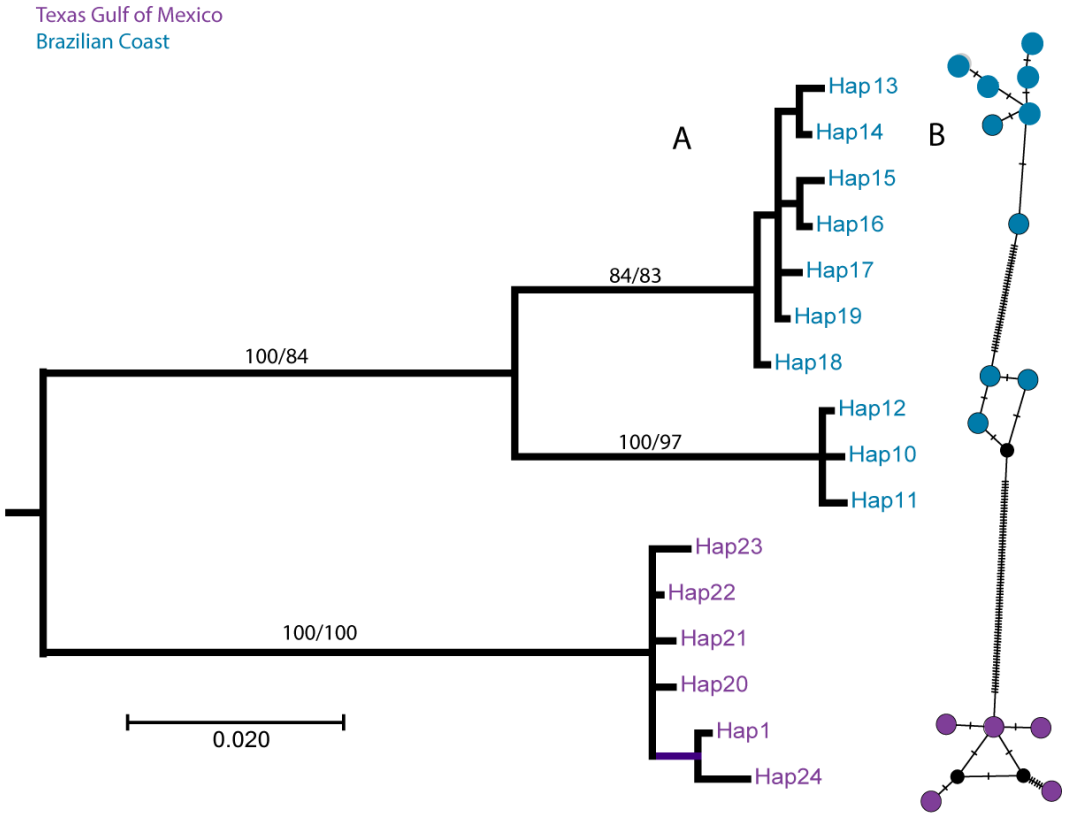

Figure 4 Phylogeographic analyses of haplogroup II. The data was divided into 2 geographic groups: Texas GOM and Brazilian coast. (**A)** Phylogeny of haplogroup I inferred by Bayesian methods. Branch values represent PP/bootstrap statistics. Tree topology based on ML methods was similar, an * was placed where bootstrap probabilities are < 0.50. Color of branches correspond to the geographic groups. (**B)** Median-joining network analysis of haplogroup I based on presence or absence of haplotypes. The size of the node does not reflect the haplotype frequency. The size of the nodes represents the number of regions where each haplotype is present. Haplotype 1 was excluded from the network analysis because it had <557 base pairs. The small, black nodes represent an inferred ancestral haplotype. Notches on each branch represent the number of nucleotide changes that occurred from one haplotype to another. The color of each node represents the location the individual was collected from (see legend).

Haplogroup III (S Clade)

The third lineage, haplogroup III, contains Clade S as defined by Chen and Hare (2008). Haplogroup III has 16 haplotypes from the Texas GOM (Bahia Grande, Port Aransas, and Galveston Bay) and 23 haplotypes from the northeastern Atlantic coast. The basal lineage of haplogroup III is a single haplotype from the northeastern Atlantic coast. The relationships within haplogroup III are largely unresolved.

While the phylogeny revealed only one clear clade, the median joining network analysis shows three clear groups. In the network analysis, the first group has 17 haplotypes from the northeastern Atlantic coast and was previously defined as the S Clade (Chen and Hare, 2008); haplotype 60 also includes an individual from Europe. The second group contains one haplotype from the northeastern Atlantic coast, three haplotypes from the Eastern GOM and one haplotype from the northeastern Atlantic coast, and 12 haplotypes from Texas, with haplotype 55 sharing individuals from the Texas GOM and the eastern GOM (Florida side). The third group has three haplotypes from the Texas GOM; haplotype 40 has two individuals, one from the Texas GOM and one from the northeastern Atlantic coast.

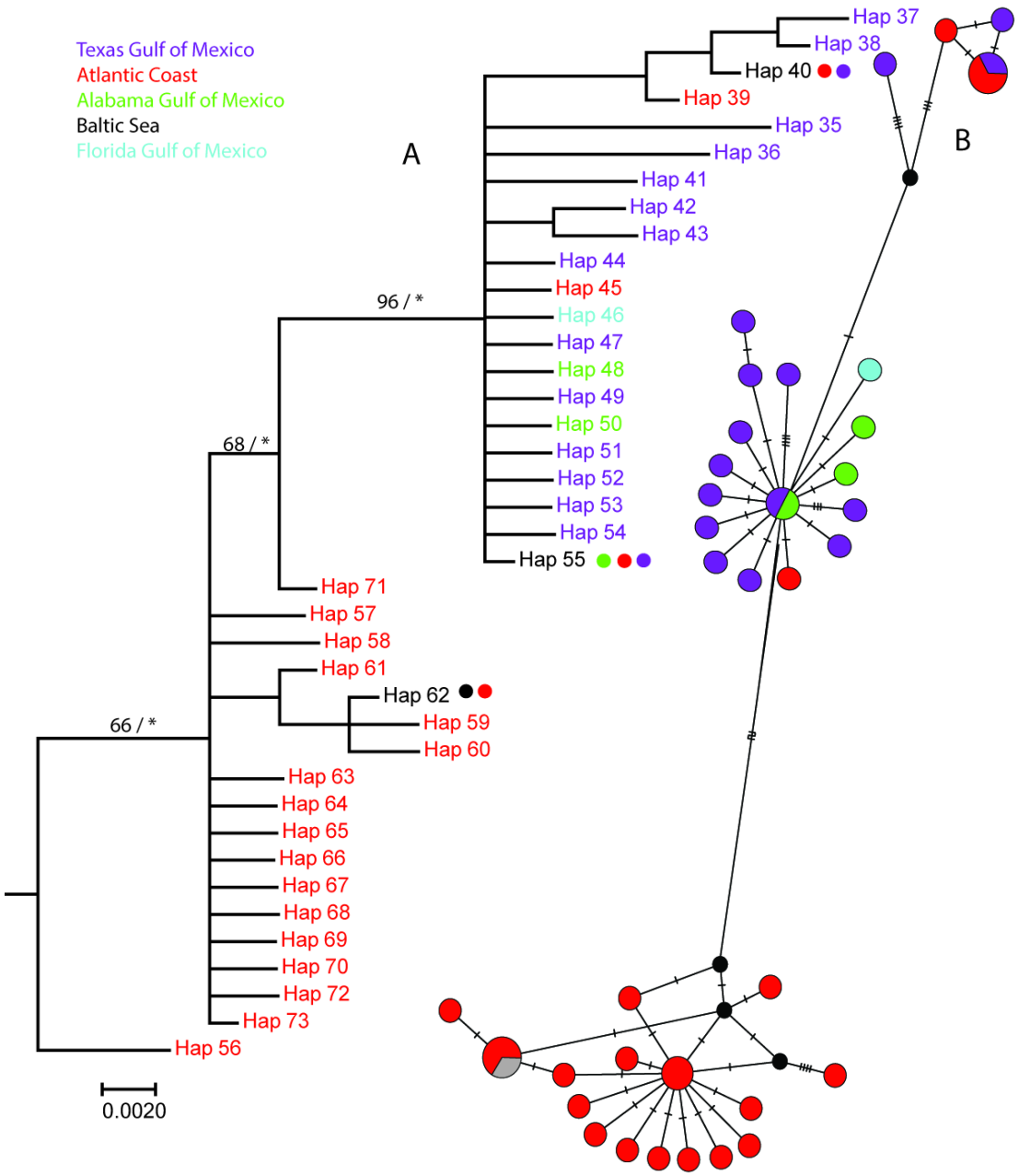

Figure 5 Phylogeographic analyses of the haplogroup III. Representatives of each haplotype were isolated from 5 geographic areas: Texas GOM, northeastern Atlantic coast, Alabama GOM, Baltic Sea and Florida GOM. (**A)** Phylogeny of haplogroup I inferred by Bayesian methods. Branch values represent PP/bootstrap statistics. Tree topology based on ML methods was similar, an * was placed where bootstrap probabilities are < 0.50. Color of branches correspond to the geographic groups. (**B)** Median-joining network analysis of haplogroup I based on presence or absence of haplotypes. The size of the node does not reflect the haplotype frequency. The size of the nodes represents the number of regions where each haplotype is present. The small, black nodes represent an inferred ancestral haplotype. Notches on each branch represent the number of changes nucleotide that occurred from one haplotype to another. The color of each node represents the location the individual was collected from (see legend).

Haplogroup IV

The fourth lineage, haplogroup IV, includes three clades. The basal clade consists of seven haplotypes from Canada and New Jersey. Sister to the basal clade there are two lineages, one corresponds to clade L3 as defined by Costa et al. (2011 and 2014) with three haplotypes from the Brazilian coast and two haplotypes from Baffin Bay the Texas GOM. The third clade contains unique haplotypes from Baffin Bay, Texas GOM. Similarly, the network analysis reveals three distinct groups within haplogroup IV. The first contains three haplotypes from the northeastern Atlantic coast. The second consists of three haplotypes from the Brazilian coast and the third has three haplotypes from the Texas GOM.

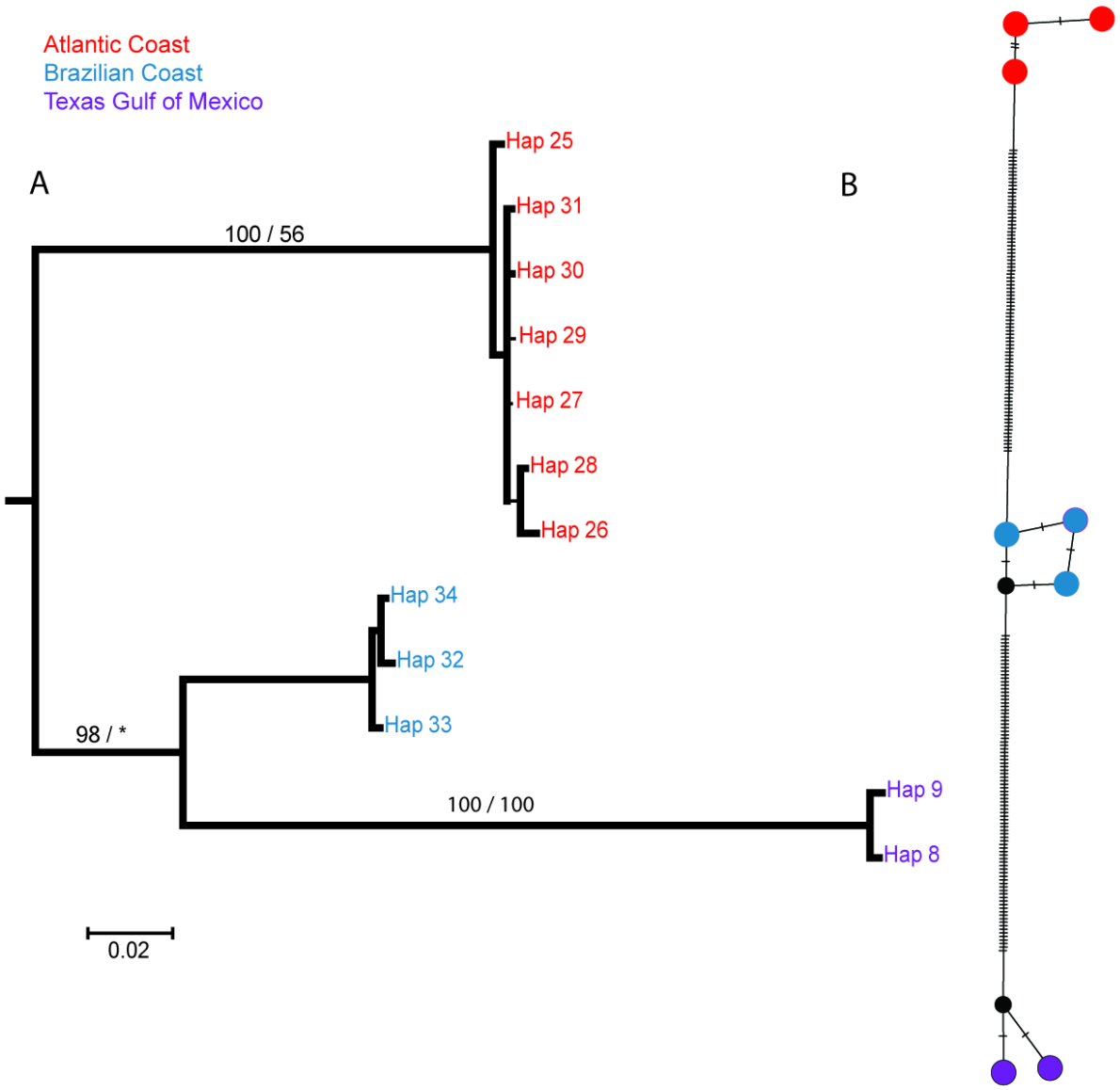

Figure 6 Phylogeographic analyses of the haplogroup IV. The data was divided into 3 geographic groups: Texas GOM, Brazilian coast, and northeastern Atlantic coast. (**A)** Phylogeny of haplogroup I inferred by Bayesian methods. Branch values represent PP/bootstrap statistics. Tree topology based on ML methods was similar, an * was placed where bootstrap probabilities are < 0.50. Color of branches correspond to the geographic groups. (**B)** Median-joining network analysis of haplogroup I based on presence or absence of haplotypes. The size of the node does not reflect the haplotype frequency. The size of the nodes represents the number of regions where each haplotype is present. Haplotypes 26, 28, 30, and 31 was excluded from the network analysis because it had <557 base pairs. The small, black nodes represent an inferred ancestral haplotype. Notches on each branch represent the number of nucleotide changes that occurred from one haplotype to another. The color of each node represents the location the individual was collected from (see legend).

Haplogroup V (F Clade)

The fifth diverging lineage, haplogroup V (Figure 7) contains the F clade, defined by Chen and Hare (2008 and 2011). Seven major lineages emerged within this clade; F1 through F5 were described by Chen and Hare (2008 and 2011) whereas lineages F6 and F7 are newly identified basal lineages in this group. Samples from the Texas GOM belong to the three basal clades (F5, F6, and F7). The basal lineage, F7 splits into six haplotypes from the Texas GOM including; five from Port Mansfield, two from Bahia Grande, six from Baffin Bay, three from Port Aransas, and two haplotypes from the Florida GOM. The second branch splits into two main lineages; one contains F5 and F6 and the other contains F1 through F4. The newly defined clade F6 comprises haplotypes from the Texas GOM including ten haplotypes from Port Lavaca and two from Galveston Bay. Clade F5 includes three haplotypes from the Texas GOM; three individuals from Bahia Grande and one individual from Port Lavaca. Clade F5 also consists of four haplotypes from the South Florida east Atlantic coast, and one haplotype from Georgia and the North Florida East Coast. The next major split contains two well-supported lineages; clade F4 and clades F1 through F3. Clade F4 branches into five haplotypes from South Florida East Coast, seven haplotypes from Georgia and the North Florida East Coast, and one haplotype from the North East Atlantic Coast. The last major lineage comprising clades F1 through F3 is not well resolved. Because the major branch is not strongly supported (posterior probability 0.94), the relationships between clades F1-F3 are not well defined. All 44 haplotypes from clades F1-F3 are from the northeastern Atlantic coast. The haplotype clusters from the median joining network analysis were consistent with the phylogeny for haplogroup V.

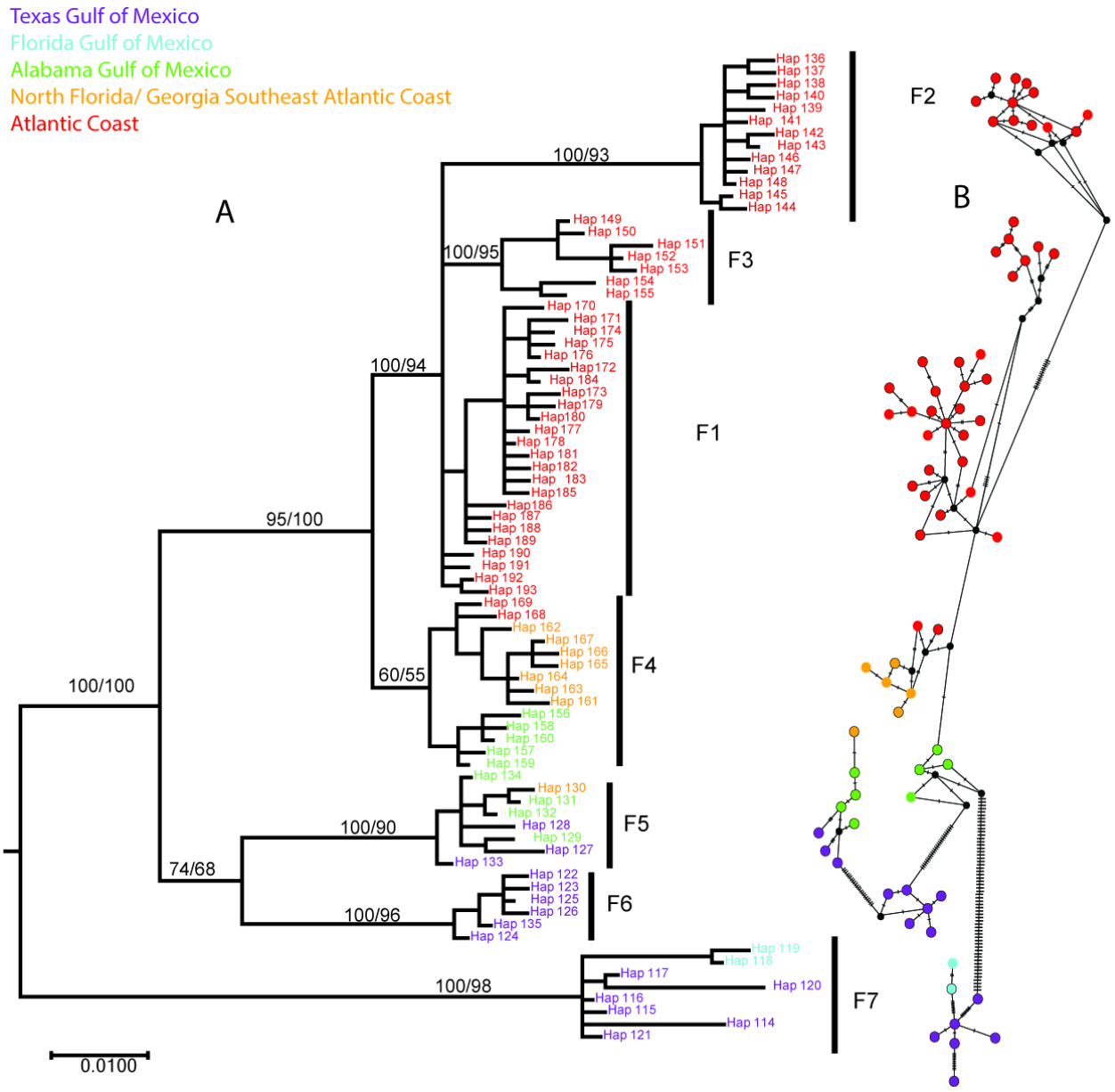

Figure 7 Phylogeographic analyses of haplogroup V. Data was divided into 5 geographic groups: Texas GOM, Florida Gulf coast, Alabama Gulf coast, north Florida/Georgia southeast Atlantic coast, and northeast Atlantic coast. (**A)** Phylogeny of haplogroup I inferred by Bayesian methods. Branch values represent PP/bootstrap statistics. Tree topology based on ML methods was similar, an * was placed where bootstrap probabilities are < 0.50. Color of branches correspond to the geographic groups. (**B)** Median-joining network analysis of haplogroup I based on presence or absence of haplotypes. The size of the node does not reflect the haplotype frequency. The small, black nodes represent an inferred ancestral haplotype. Notches on each branch represent the number of nucleotide changes that occurred from one haplotype to another. The color of each node represents the location the individual was collected from (see legend).

Haplogroup VI (X Clade)

The sixth diverging lineage, haplogroup VI corresponds to the X Clade defined by Chen and Hare (2011), which includes 40 haplotypes from the North East Atlantic Coast and Europe. No specimens from the Texas GOM belong to this clade.

Texas GOM Network Analysis

The median-joining network with the samples sequenced from the Texas GOM shows seven distinct groups that correspond to haplogroups I-V. For this network analysis, the size of the nodes is representative of the haplotype frequencies for each haplotype. In haplogroup I, one haplotype was collected in Port Aransas. In haplogroup II five haplotypes originate from Port Mansfield. Haplogroup III contains one haplotype with individuals from Port Aransas, Bahia Grande and Galveston Bay, one haplotype with individuals from Bahia Grande and Port Aransas, four haplotypes from Port Aransas, and seven haplotypes from Galveston Bay. Haplogroup IV includes two representatives from Baffin Bay. Haplogroup five is divided further into three groups; group F5, which has three haplotypes from Bahia Grande group F6, which has one haplotype with individuals from Port Mansfield and Galveston Bay and four haplotypes from Port Lavaca, and group F7, which is the most diverse. F7 has one haplotype consisting of individuals from Baffin Bay, Bahia Grande, Port Mansfield, and Port Aransas. Group F7 also has a haplotype with individuals from Port Aransas and Port Mansfield. In addition, one haplotype was found at each of the following sites; Bahia Grande, Baffin Bay, Port Mansfield, and Port Lavaca.

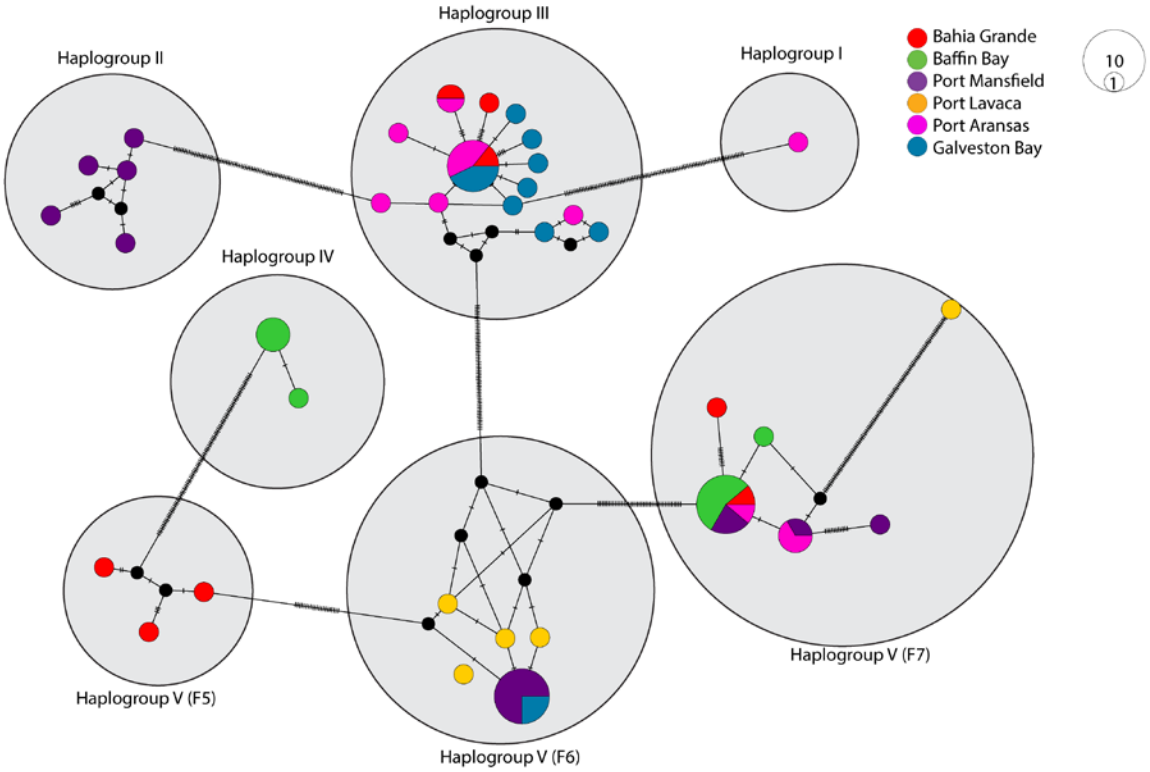

Figure 8 Median joining network analysis of the haplotypes found in the Texas GOM. Seven major groups are indicated by gray circles with their corresponding haplogroup number. The size of the node reflects the haplotype frequency. Black nodes represent an inferred ancestor. Notches on each branch represent the number of changes that occurred from one haplotype to another. The color of each node represents the location that the individual was collected (see legend).

Table 2 Percent divergence between haplogroups based on a 557 bp region of mtCOI. Calculated by summing the model-corrected branch lengths from Bayesian analysis.

	I	II	III	IV	V
II	16.4				
III	20.9	19.5			
IV	21.2	19.8	14.7		
V	26.0	24.6	19.5	12.5	
VI	28.4	27.0	21.9	14.9	13.8

CHAPTER IV

DISCUSSION

Range Expansion and Cryptic Speciation Follows the Prevailing Currents

Results support the hypothesis that southern lineages are basal to northeastern lineages (within Haplogroup V), suggesting a range expansion and species divergence that followed the prevailing circulation pattern in the GOM and north Atlantic in this haplogroup (Figure 7). This pattern is consistent with the well-known pattern called phylogeographical structuring (Soltis et al., 2006), which says that biota presently occupying regions north of the extent of the last glacial period in the Pleistocene are derived from older lineages that occupy more southern latitudes that are south of the glacial extent. It is important to note that phylogeographical structuring does not typically occur in species that have the capacity to migrate. For example, phylogeographical structuring has not been shown in plants, insects, or mammals (Soltis et al., 2006). The glacial refugia theory states that as glaciers grow, flora and fauna are displaced (Soltis et al., 2006). The patterns that are commonly seen in estuarian species such as *A. tonsa* and may be due to the isolation of populations during the last glacial cycle in the Pleistocene epoch (Jacobson et al., 1987; Soltis et al., 2006). Phylogeographical structuring and the glacial refugia theory are both patterns of divergence are consistent with my findings of the northward diversification of *A. tonsa*.

The pattern of northward expansion and speciation observed in *A. tonsa* is similar to patterns described in ctenophores (Bayha et al., 2015), amphipods (Kelly et al., 2015) black tip sharks (Keeney, 2004), barnacles (Govindarajan et al., 2015), squid (Herke and Foltz, 2001) and many

other taxa, many of which are summarized in Soltis et al., (2006). Bayha et al (2015) looked at the worldwide phylogeographic patterns of *M. leidyi* using cytochrome oxidase b (cytb) and six micronuclear satellites. When looking at cytb and six micronuclear satellites, both markers showed a separation in the species at the Labrador current suggesting that endemic lineages of *M. leidyi* from North America's Atlantic coast are kept separate by the collision of the warm Gulf Current flowing north and the cooler Labrador Current flowing south. The collision of the Gulf and Labrador currents could be an additional explanation as to why newer lineages of *A. tonsa* are only found in the northeastern Atlantic coast and why no representatives from those newer lineages are present in the Texas GOM. Basal lineages of *M. leidyi* were found in the Florida GOM while newer lineages were identified in Cape Hatteras, North Carolina. Cape Hatteras is described as genetic break for many species including fish (Avise, 1987), crustaceans (Kelly et al., 2006; Chen and Hare, 2011), and mollusks (Baker et al., 2008). The genetic break commonly described at Cape Hatteras is likely due to organism's inability to swim against the currents or their intolerance to thermal gradients (Bayha et al., 2015).

Within haplogroup V (F Clade), a break was found in the Cape Hatteras region with haplogroup V's F1-F3 lineages only being found near Chesapeake Bay and the F4 lineage residing in the Carolinas, which is south of Cape Hatteras (Chen and Hare, 2011). The F5-F7 clades have only been found south of the Atlantic coast of north Florida and the GOM.

A post-Pleistocene divergence pattern like the one observed in this study was described in a widespread amphipod, *Gammarus tigrinus* (Kelly et al., 2015). Kelley et al. (2015) showed a divergence between southern and northern lineages of *G. tigrinus* in the mid-Atlantic like the pattern found when examining *M. leidyi* (Bayha et al. 2015). The divergence of northern and southern lineages of *G. tigrinus* was supported with congruence of COI and ITSI sequencing. MtCOI

sequencing (Govindarajan et al. 2015) revealed a distinct pattern of speciation in the wide-spread barnacle, *Chthamalus fragilis*, between southern and northern sites ranging from Tampa Bay, Florida to Cape Cod, Massachusetts.

Connectivity from Brazil to the Northeastern Atlantic Coast

Despite having clearly defined clades within *A. tonsa*, connectivity within lineages of Haplogroup IV appears to span multiple regions and thousands of kilometers, linking Brazil to the GOM and the GOM with the northeast Atlantic coast. There are two clades from Brazil in Haplogroup II are sisters to a clade with haplotypes unique to the Texas site in Port Mansfield. This suggests that divergence is occurring as copepods are carried by the South Equatorial Current from coastal Brazil into the GOM. In Haplogroup IV, a common ancestor is shared with three haplotypes from Brazil and two haplotypes from Baffin Bay, Texas. These clades also share a common ancestor to haplotypes found in New Bedford, New Hampshire. This dataset is not sufficiently robust enough to define clear patterns, but does allow a conclusion that connectivity existed sometime in the past within these lineages. Because of the circulation patterns of the South Equatorial Current and the Gulf Stream and *A. tonsa's* being a holoplanktonic species, a basal clade that originated from Brazil would be expected to be the source. More sampling between Brazil and Texas may provide further answers about the evolutionary processes within these haplogroups.

Divergence Between Haplogroups

Genetic variation of mtCOI between species has been shown useful to distinguish closely related species from one other (Bucklin et al., 2002). The mtCOI sequence divergence between species of calanoid copepods ranged from 9-25% (Bucklin et al. 1998, 1999, 2001; Hill et al., 2001). Bucklin et al. (2002) sequenced mtCOI for 34 calanoid copepod species and found that

mtCOI sequence divergence between genera ranged from 12 to 25%. Congeneric species that are very similar or morphologically indistinguishable (cryptic species) can display significant mitochondrial genes sequence divergence (Hill and Bucklin, 2001). For example, mtCOI sequence divergence ranged between 8-12% among 18 calanoid copepod species of six genera (Bucklin et al., 1999). My data are consistent with the upper range of the divergences found for interspecific and between different genera divergence for calanoid copepods with divergence between Haplogroups I-VI ranging between 12.5-28.4%. Hirai et al. (2015) found 18 divergent lineages of *Pleuromamma abdominalis* that they hypothesized to be 18 distinct species based on percent divergences ranging from 4.3-28.6%. In addition to the high genetic distance between the 18 lineages of *P. abdominalis,* Hirai et al. (2015) also noted some morphological differentiation among the lineages.

Because divergences among haplotypes in this study are within or above the range described for congeneric divergences, each haplogroup should be labeled as distinct species based on the evolutionary species concept; where species can be identified based on their level of genetic differences (Hill and Bucklin, 2001). Based on the mitochondrial RNA subunit 16S, divergence between three *Acartia* species ranged from 19-28% (Bucklin et al., 2002). Congeneric species have been reported having up to 25% differences in their sequences (Bucklin et al., 2002).

The Texas GOM Phylogenetics and Salinity

The results show here indicate that gene flow exists between Bahia Grande, Baffin Bay, Port Lavaca, and Port Mansfield within the oldest lineage (F7) in haplogroup V (Figure 7.0). The appearance of numerous haplotypes in these areas may result from the heavy use of these bays as shipping ports. In some instances, transport occurring via ballast water could be responsible for spreading copepods beyond their natural expansive range. In 2005 the Bahia Grande restoration

project began and continues to restore the estuary to its historical flooding patterns and salinities with the goal being to provide wetlands for ecological diversity, mainly birds (Marquez et al., 2016). Water was flooded back into Bahia Grande via the Brownsville Shipping Channel on the north side and a restoration channel on the south side (Marquez et al., 2016). The shipping channel is a heavily traveled commercial shipping lane. The re-flooding of Bahia Grande via the shipping and restoration channels (connected to the GOM) are most likely responsible for facilitating spread of the various *A. tonsa* haplotypes found at that site. Additional sampling in the Gulf side (Northern Bahia Grande) and the ocean side of South Padre Island may provide answers as to where the haplotypes found in Bahia Grande are coming from.

1) Haplogroup V

Within haplogroup V (Figure 7), a clear pattern of speciation is occurring as the species made a northward expansion since the last glacial cycle during the Pleistocene epoch. The lineages from the Texas and Florida GOM are the oldest lineages that have been described within the F Clade. This basal clade, F7, shows a connection between the eastern and western GOM, probably due to a historic connection that was not affected by the last glacial period. The latitude where the sampling for this project took place did not experience geographical changes due to climactic changes during the last glacial period. Because of this, *A. tonsa* was not displaced as they were in the Northern Atlantic estuaries (Chen and Hare, 2008 & 2011).

2) Haplogroup III

Representatives of haplogroup III (S Clade) (Figure 5) in the Texas GOM do not display the same geographical pattern of connectivity as observed in the other haplogroups. Two well defined clades were identified within the haplogroup III. The most basal lineage is a haplotype from the northeastern Atlantic Coast. In the second subclade, relationships between haplotypes

cannot be resolved due to low posterior probabilities and bootstrap probabilities of branch support. Chen and Hare (2011) proposed that the lack of resolution in their S clade was likely due to higher migration rates because of higher salinity tolerances. Larger sizes and additional genetic markers may resolve these relationships and a similar pattern of divergence may be observed.

3) Salinity

The F clade was previously described by Chen and Hare (2008) as having an affinity for lower salinities (0.3-12 psu). Lineages that are basal to the F clade have now been found in the Texas GOM which has salinity ranging from 2.6-32.4 psu (during the time of field sampling for this project) (Table 1). Some of the sampling sites for this project are within the largest hypersaline system in the world (Marquez et al., 2016). This finding suggests that the F clade originated from waters that are hypersaline in some cases, and that the divergence and affinity to lower saline waters is a more recent one that only occurs in the northeast Atlantic estuaries.

The S clade was described as having an affinity to a broad range of salinities and individuals from this clade were found in salinities ranging from 2-26 psu in the Northeastern Atlantic Coast (Chen and Hare, 2008). Haplotypes from the Texas GOM belonging to a sister clade to the S clade were collected in salinities ranging from 2.6-32.4 psu (Table 1). The highest salinity measured in the Texas GOM is 6.4 psu higher than that measured in Chesapeake Bay by Chen and Hare. This finding shows that the S clade and its sister clade are both tolerant of higher salinity in the Texas GOM.

Future Research

Further sampling in South America, the Caribbean Mexico and along the Pacific coastal estuaries may provide additional resolution for understanding phylogeographical patterns of *A. tonsa*. Future studies should use multiple genetic markers in addition to mtCOI as there are many

unresolved relationships within the S Clade that would likely be resolved. Also, phenotypic analysis should be done on all the lineages including identifying any morphological characteristics that may differ between the lineages. Finally, observing behavioral characteristics and performing experiments on biochemical properties for the different lineages could be done.

Broader Impacts

This study begins to provide a holistic picture of the evolutionary history of *A. tonsa* by-presenting a phylogeographic analysis of *A. tonsa* from the Texas GOM. Examination of the genetic diversity of *A. tonsa* in the Texas GOM contributes to improve our understanding of the patterns of speciation that occur in other marine species due to global climactic changes such as the last glacial period during the Pleistocene epoch. We now also have information about lineages of *A. tonsa* that inhabit the Texas GOM bays, and this knowledge will be submitted to GenBank to become accessible to researchers.

REFERENCES

Avise, J. C., Reeb, C. A., & Saunders, N. C. (1987). Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). *Evolution*, 991-1002.

Avise, J. C. (2000). *Phylogeography: the history and formation of species*. Harvard university press.

Baker, P., Austin, J. D., Bowen, B. W., & Baker, S. M. (2008). Range-wide population structure and history of the northern quahog (Merceneria merceneria) inferred from mitochondrial DNA sequence data. *ICES Journal of Marine Science, 65*(2), 155-163.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., ... & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in ecology & evolution*, *22*(3), 148-155.

Blaxter, J. H., Douglas, B., Tyler, P. A., & Mauchline, J. (1998). *The biology of calanoid copepods* (Vol. 33). Academic Press.

Bucklin, A., La Jeunesse, T. C., Curry, E., Wallinga, J., & Garrison, K. (1996). Molecular diversity of the copepod, Nannocalanus minor: genetic evidence of species and population structure in the North Atlantic Ocean. *Journal of Marine Research*, *54*(2), 285-310.

Bucklin, A., Guarnieri, M., Hill, R. S., Bentley, A. M., & Kaartvedt, S. (1999). Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. *Hydrobiologia*, *401*, 239-254.

Bucklin, A., Frost, B., Bradford-Grieve, J., Allen, L., & Copley, N. (2003). Molecular systematic and phylogenetic assessment of 34 calanoid copepod species of the Calanidae and Clausocalanidae. *Marine Biology*, *142*(2), 333-343.

Caudill, C. C., & Bucklin, A. (2004). Molecular phylogeography and evolutionary history of the estuarine copepod, Acartia tonsa, on the Northwest Atlantic coast. *Hydrobiologia*, *511*(1), 91-102.

Chen, G., & Hare, M. P. (2008). Cryptic ecological diversification of a planktonic estuarine copepod, Acartia tonsa. *Molecular Ecology*, *17*(6), 1451-146

Chen, G., & HARE, M. P. (2011). Cryptic diversity and comparative phylogeography of the estuarine copepod Acartia tonsa on the US Atlantic coast. *Molecular Ecology*, *20*(11), 2425-2441.

Costa, K. G., Rodrigues Filho, L. F. S., Costa, R. M., Vallinoto, M., Schneider, H., & Sampaio, I. (2014). Genetic variability of Acartia tonsa (Crustacea: Copepoda) on the Brazilian coast. *Journal of Plankton Research*, *36*(6), 1419-1422.

Dyke, Arthur, and Victor Prest. "Late Wisconsinan and Holocene history of the Laurentide ice sheet." *Géographie physique et Quaternaire* 41.2 (1987): 237-263.

Edgar, Robert C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic acids research* 32.5 (2004): 1792-1797.

Eunmi Lee, C. (2000). Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". *Evolution*, *54*(6), 2014-2027.

Gillespie, M. C. (1971). Analysis and treatment of zooplankton of estuarine waters of Louisiana.

Goetze, E. (2003). Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*(1531), 2321-2331.

Gonzalez, J. G. (1974). Critical thermal maxima and upper lethal temperatures for the calanoid copepods Acartia tonsa and A. clausi. *Marine Biology*, *27*(3), 219-223.

Hill, R., Allen, L., & Bucklin, A. (2001). Multiplexed species-specific PCR protocol to discriminate four N. Atlantic Calanus species, with an mtCOI gene tree for ten Calanus species. *Marine Biology*, *139*(2), 279-287.

Hirai, J., Tsuda, A., & Goetze, E. (2015). Extensive genetic diversity and endemism across the global range of the oceanic copepod Pleuromamma abdominalis. *Progress in Oceanography*, *138*, 77-90.

Huelsenbeck, J. P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17:754-755.

Jacobson Jr, G. L., Webb III, T., & Grimm, E. C. (1987). Patterns and rates of vegetation change during the deglaciation of eastern North America. *North America and adjacent oceans during the last deglaciation*, *3*, 277-288.

Kelly, D. W., MacIsaac, H. J., & Heath, D. D. (2006). Vicariance and dispersal effects on phylogeographic structure and speciation in a widespread estuarine invertebrate. *Evolution*, *60*(2), 257-267.

Lance, J. (1964). The salinity tolerances of some estuarine planktonic crustaceans. *The Biological Bulletin*, *127*(1), 108-118.

Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110-1116.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451-1452.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29(6):1695-1701.http://dx.doi.org/10.1093/molbev/mss020

Magalhães, A., Nobre, D. S. B., Bessa, R. S. C., Pereira, L. C. C., & Da Costa, R. M. (2011). Seasonal and short-term variations in the copepod community of a shallow Amazon estuary (Taperaçu, Northern Brazil). *Journal of Coastal Research*, (64), 1520.

Marquez, M. A., Fierro-Cabo, A., & Cintra-Buenrostro, C. E. (2016). Can ecosystem functional recovery be traced to decomposition and nitrogen dynamics in estuaries of the Lower Laguna Madre, Texas? *Restoration Ecology*.

Peijnenburg, K. T. C. A., Breeuwer, J. A., Pierrot-Bults, A. C., & Menken, S. B. (2004). Phylogeography of the planktonic chaetognath Sagitta setosa reveals isolation in European seas. *Evolution*, *58*(7), 1472-1487.

Sáez, A. G., Probert, I., Geisen, M., Quinn, P., Young, J. R., & Medlin, L. K. (2003). Pseudo-cryptic speciation in coccolithophores. *Proceedings of the National Academy of Sciences*, *100*(12), 7163-7168.

Stamatakis, Alexandros. "RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies." *Bioinformatics* 30.9 (2014): 1312–1313. *PMC*. Web. 30 Nov. 2017.

Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol*, *3*(5), 294-299.

Zachos, F. E. (2016). *Species Concepts in Biology*. Springer International Pu.

Zane, L., Ostellari, L., Maccatrozzo, L., Bargelloni, L., Cuzin-Roudy, J., Buchholz, F., & Patarnello, T. (2000). Genetic differentiation in a pelagic crustacean (Meganyctiphanes norvegica: Euphausiacea) from the North East Atlantic and the Mediterranean Sea. *Marine Biology*, *136*(2), 191-199.

APPENDIX

Appendix 1 Haplotype numbers with their corresponding GenBank accession numbers and sample names for the samples collected in the Texas GOM.

GenBank Accession	Haplotype Number	Haplogroup	Location	Sample Number
Present study	1	2	Port Mansfield	PA-10
KC287392	2	1	Rhode Island	
KC287393	3	1	Rhode Island	
KC287394	3	1	Rhode Island	
KC287395	3	1	Rhode Island	
KC287397	3	1	Rhode Island	
KC287396	4	1	Rhode Island	
KC287398	5	1	Rhode Island	
Present study	6	1	Port Aransas	PA-7
KC287386	7	1	Southern California Coast	
KC287387	7	1	Southern California Coast	
Present Study	8	4	Baffin Bay, Texas	BB-2
Present Study	9	4	Baffin Bay, Texas	BB-1
Present Study	9	4	Baffin Bay, Texas	BB-8
Present Study	9	4	Baffin Bay, Texas	BB-3
KM458086	10	2	Brazil	
KM458087	11	2	Brazil	
KM458085	12	2	Brazil	
KM458083	13	2	Brazil	
KM458082	14	2	Brazil	
KM458078	15	2	Brazil	
KM458079	16	2	Brazil	
KM458080	17	2	Brazil	
KM458084	18	2	Brazil	

KM458081	19	2	Brazil	
Present Study	20	2	Port Mansfield, TX	PM 5-1
Present Study	21	2	Port Mansfield, TX	PM 10-1
Present Study	22	2	Port Mansfield, TX	PM 3-1
Present Study	23	2	Port Mansfield, TX	PM 4-1
Present Study	24	2	Port Mansfield, TX	PM 2-1
KC287407	25	4	New Bedford, MA	
KC287403	26	4	New Bedford, MA	
KC287405	27	4	New Bedford, MA	
KC287408	27	4	New Bedford, MA	
KC287402	28	4	New Bedford, MA	
KC287404	29	4	New Bedford, MA	
KC287399	30	4	New Bedford, MA	
KC287400	30	4	New Bedford, MA	
KC287401	31	4	New Bedford, MA	
KM458076	32	4	Brazil	
KM458077	33	4	Brazil	
KM458075	34	4	Brazil	
Present Study	35	3(S)	Bahia Grande, TX	BG 1
Present Study	36	3(S)	Galveston Bay, TX	KB 5
Present Study	37	3(S)	Galveston Bay, TX	KB 7
Present Study	38	3(S)	Port Aransas, TX	PA 8
JF304055	39	3(S)	Atlantic Coast	
JF304029	40	3(S)	Atlantic Coast	
Present Study	40	3(S)	Galveston Bay, TX	KB-11
Present Study	41	3(S)	Bahia Grande, TX	BG5
Present Study	42	3(S)	Galveston Bay, TX	KB 10
Present Study	43	3(S)	Port Aransas, TX	PA 8

Present Study	44	3(S)	Bahia Grande, TX	Ship Channel 12
JF304080	45	3(S)	Atlantic Coast	
DQ431912	46	3(S)	Florida	
Present Study	47	3(S)	Port Aransas, TX	PA 10
EU196724	48	3(S)	Mobile Bay, AL	
DQ431908	50	3(S)	Mobile Bay, AL	
Present Study	51	3(S)	Port Aransas, TX	PA 11
Present Study	52	3(S)	Galveston Bay, TX	KB 12
Present Study	53	3(S)	Galveston Bay, TX	KB 8
Present Study	54	3(S)	Galveston Bay, TX	KB 9
EU196720	55	3(S)	Mobile Bay, AL	
KC287416	55	3(S)	Pamlico Sound, NC	
JF304051	55	3(S)	Atlantic Coast	
Present Study	55	3(S)	Port Aransas, TX	PA 1
Present Study	55	3(S)	Port Aransas, TX	PA 4
Present Study	55	3(S)	Galveston Bay, TX	KB 4
Present Study	55	3(S)	Port Aransas, TX	PA 12
Present Study	55	3(S)	Galveston Bay, TX	KB 6
EU196719	55	3(S)	Mobile Bay, AL	
EU196721	55	3(S)	Mobile Bay, AL	
EU196722	55	3(S)	Mobile Bay, AL	
EU196723	55	3(S)	Mobile Bay, AL	
DQ431907	55	3(S)	Mobile Bay, AL	
DQ431911	56	3(S)	Turkey Point, FL	
KC287421	56	3(S)	Rhode Island	
JF304065	57	3(S)	Atlantic Coast	
EU274442	58	3(S)	Atlantic Coast	
JF304069	59	3(S)	Atlantic Coast	

JF304074	60	3(S)	Atlantic Coast	
EU274438	61	3(S)	Atlantic Coast	
EU274440	62	3(S)	Atlantic Coast	
KC287385	62	3(S)	Atlantic Coast	
KC287406	62	3(S)	Atlantic Coast	
KC287409	62	3(S)	Atlantic Coast	
JF304028	62	3(S)	Baltic Sea	
EU016219	63	3(S)	Atlantic Coast	
JF304093	64	3(S)	Atlantic Coast	
JF304070	65	3(S)	Atlantic Coast	
JF304072	66	3(S)	Atlantic Coast	
JF304073	67	3(S)	Atlantic Coast	
JF304092	68	3(S)	Atlantic Coast	
JF304094	69	3(S)	Atlantic Coast	
EU274437	70	3(S)	Atlantic Coast	
EU274439	71	3(S)	Atlantic Coast	
EU274441	72	3(S)	Atlantic Coast	
EU274443	73	3(S)	Atlantic Coast	
EU274436	74	6(X)	Canadian Coast	
GQ466417	75	6(X)	Damariscotta, ME	
KC287388	75	6(X)	Damariscotta, ME	
KC287411	76	6(X)	Atlantic Coast	
KC287417	77	6(X)	Atlantic Coast	
JF304034	78	6(X)	Atlantic Coast	
JX995268	79	6(X)	Germany Coast	
JX995260	80	6(X)	Germany Coast	
JX995270	80	6(X)	Germany Coast	
JF304075	81	6(X)	Atlantic Coast	

KC287288	82	6(X)	Atlantic Coast	
KC287310	83	6(X)	Atlantic Coast	
KC287273	83	6(X)	Atlantic Coast	
KC287274	83	6(X)	Atlantic Coast	
KC287275	83	6(X)	Atlantic Coast	
KC287277	83	6(X)	Atlantic Coast	
KC287279	83	6(X)	Atlantic Coast	
KC287280	83	6(X)	Atlantic Coast	
KC287284	83	6(X)	Atlantic Coast	
KC287285	83	6(X)	Atlantic Coast	
KC287286	83	6(X)	Atlantic Coast	
KC287287	83	6(X)	Atlantic Coast	
KC287290	83	6(X)	Atlantic Coast	
KC287291	83	6(X)	Atlantic Coast	
KC287294	83	6(X)	Atlantic Coast	
KC287296	83	6(X)	Atlantic Coast	
KC287298	83	6(X)	Atlantic Coast	
KC287311	83	6(X)	Atlantic Coast	
KC287312	83	6(X)	Atlantic Coast	
KC287316	83	6(X)	Atlantic Coast	
KC287317	83	6(X)	Atlantic Coast	
KC287321	83	6(X)	Atlantic Coast	
KC287324	83	6(X)	Atlantic Coast	
KC287326	83	6(X)	Atlantic Coast	
KC287330	83	6(X)	Atlantic Coast	
KC287332	83	6(X)	Atlantic Coast	
KC287333	83	6(X)	Atlantic Coast	
KC287336	83	6(X)	Atlantic Coast	

KC287339	83	6(X)	Atlantic Coast	
KC287342	83	6(X)	Atlantic Coast	
KC287343	83	6(X)	Atlantic Coast	
KC287344	83	6(X)	Atlantic Coast	
KC287365	83	6(X)	Atlantic Coast	
KC287369	83	6(X)	Atlantic Coast	
KC287370	83	6(X)	Atlantic Coast	
KC287374	83	6(X)	Atlantic Coast	
KC287375	83	6(X)	Atlantic Coast	
KC287378	83	6(X)	Atlantic Coast	
KC287380	83	6(X)	Atlantic Coast	
KC287382	83	6(X)	Atlantic Coast	
KC287383	83	6(X)	Atlantic Coast	
KC287390	83	6(X)	Atlantic Coast	
KC287418	83	6(X)	Atlantic Coast	
KC287419	83	6(X)	Atlantic Coast	
JF304033	90	6(X)	Atlantic Coast	
KC287345	90	6(X)	Atlantic Coast	
KC287346	90	6(X)	Atlantic Coast	
KC287347	90	6(X)	Atlantic Coast	
KC287349	90	6(X)	Atlantic Coast	
KC287352	90	6(X)	Atlantic Coast	
KC287354	90	6(X)	Atlantic Coast	
KC287356	90	6(X)	Atlantic Coast	
KC287357	90	6(X)	Atlantic Coast	
KC287358	90	6(X)	Atlantic Coast	
KC287359	90	6(X)	Atlantic Coast	
KC287360	90	6(X)	Atlantic Coast	

KC287364	91	6(X)	Atlantic Coast	
KC287420	92	6(X)	Atlantic Coast	
KC287389	92	6(X)	Atlantic Coast	
KC287410	92	6(X)	Atlantic Coast	
KC287412	92	6(X)	Atlantic Coast	
KC287413	92	6(X)	Atlantic Coast	
KC287414	92	6(X)	Atlantic Coast	
KC287415	93	6(X)	Atlantic Coast	
JF304071	94	6(X)	Atlantic Coast	
JF304037	95	6(X)	Atlantic Coast	
KC287371	96	6(X)	Atlantic Coast	
KC287373	97	6(X)	Atlantic Coast	
KC287367	98	6(X)	Atlantic Coast	
JF304047	99	6(X)	Atlantic Coast	
JF304030	100	6(X)	Atlantic Coast	
HE647798	100	6(X)	Italy Coast	
JF304032	101	6(X)	Atlantic Coast	
KC287315	102	6(X)	Atlantic Coast	
KC287338	103	6(X)	Germany Coast	
JX995266	103	6(X)	Germany Coast	
JX995267	104	6(X)	Germany Coast	
EU196702	105	6(X)	Denmark Coast	
KC287276	105	6(X)	Atlantic Coast	
KC287281	105	6(X)	Atlantic Coast	
KC287289	105	6(X)	Atlantic Coast	
KC287292	105	6(X)	Atlantic Coast	
KC287293	105	6(X)	Atlantic Coast	
KC287295	105	6(X)	Atlantic Coast	

KC287297	105	6(X)	Atlantic Coast	
KC287302	105	6(X)	Atlantic Coast	
KC287303	105	6(X)	Atlantic Coast	
KC287304	105	6(X)	Atlantic Coast	
KC287305	105	6(X)	Atlantic Coast	
KC287307	105	6(X)	Atlantic Coast	
KC287308	105	6(X)	Atlantic Coast	
KC287309	105	6(X)	Atlantic Coast	
KC287314	105	6(X)	Atlantic Coast	
KC287320	105	6(X)	Atlantic Coast	
KC287323	105	6(X)	Atlantic Coast	
KC287325	105	6(X)	Atlantic Coast	
KC287335	105	6(X)	Atlantic Coast	
KC287337	105	6(X)	Atlantic Coast	
KC287340	105	6(X)	Atlantic Coast	
KC287366	105	6(X)	Atlantic Coast	
KC287368	105	6(X)	Atlantic Coast	
KC287372	105	6(X)	Atlantic Coast	
KC287376	105	6(X)	Atlantic Coast	
KC287377	105	6(X)	Atlantic Coast	
KC287384	105	6(X)	Atlantic Coast	
JF304077	106	6(X)	Denmark Coast	
EU196716	107	6(X)	Atlantic Coast	
JF304050	108	6(X)	Atlantic Coast	
KC287438	109	6(X)	Atlantic Coast	
KC287350	109	6(X)	Atlantic Coast	
KC287531	109	6(X)	Atlantic Coast	
KC287353	110	6(X)	Atlantic Coast	

KC287362	110	6(X)	Atlantic Coast	
KC287355	111	6(X)	Atlantic Coast	
KC287361	112	6(X)	Atlantic Coast	
KC287363	112	6(X)	Atlantic Coast	
JF304088	113	6(X)	Atlantic Coast	
KC287272	113	6(X)	Atlantic Coast	
KC287278	113	6(X)	Atlantic Coast	
KC287282	113	6(X)	Atlantic Coast	
KC287283	113	6(X)	Atlantic Coast	
KC287299	113	6(X)	Atlantic Coast	
KC287300	113	6(X)	Atlantic Coast	
KC287301	113	6(X)	Atlantic Coast	
KC287306	113	6(X)	Atlantic Coast	
KC287313	113	6(X)	Atlantic Coast	
KC287318	113	6(X)	Atlantic Coast	
KC287319	113	6(X)	Atlantic Coast	
KC287322	113	6(X)	Atlantic Coast	
KC287327	113	6(X)	Atlantic Coast	
KC287328	113	6(X)	Atlantic Coast	
KC287329	113	6(X)	Atlantic Coast	
KC287331	113	6(X)	Atlantic Coast	
KC287334	113	6(X)	Atlantic Coast	
KC287341	113	6(X)	Atlantic Coast	
KC287348	113	6(X)	Atlantic Coast	
KC287350	113	6(X)	Atlantic Coast	
KC287351	113	6(X)	Atlantic Coast	
KC287353	113	6(X)	Atlantic Coast	
KC287355	113	6(X)	Atlantic Coast	

KC287361	113	6(X)	Atlantic Coast	
KC287362	113	6(X)	Atlantic Coast	
KC287363	113	6(X)	Atlantic Coast	
KC287379	113	6(X)	Atlantic Coast	
KC287381	113	6(X)	Atlantic Coast	
KC287391	113	6(X)	Atlantic Coast	
JX995255	113	6(X)	Germany Coast	
JX995256	113	6(X)	Germany Coast	
JX995257	113	6(X)	Germany Coast	
JX995258	113	6(X)	Germany Coast	
JX995259	113	6(X)	Germany Coast	
JX995261	113	6(X)	Denmark Coast	
JX995262	113	6(X)	Denmark Coast	
JX995263	113	6(X)	Denmark Coast	
JX995264	113	6(X)	Denmark Coast	
JX995265	113	6(X)	Germany Coast	
JX995269	113	6(X)	Germany Coast	
HE647797	113	6(X)	Germany Coast	
JF304027	113	6(X)	Atlantic Coast	
EU196703	113	6(X)	Germany Coast	
EU196704	113	6(X)	Germany Coast	
EU196705	113	6(X)	Germany Coast	
EU196706	113	6(X)	Germany Coast	
EU196707	113	6(X)	Germany Coast	
EU196708	113	6(X)	Germany Coast	
EU196709	113	6(X)	Germany Coast	
EU196710	113	6(X)	Germany Coast	
EU196711	113	6(X)	Germany Coast	

EU196712	113	6(X)	Denmark Coast	
EU196713	113	6(X)	Denmark Coast	
EU196714	113	6(X)	Denmark Coast	
EU196715	113	6(X)	Denmark Coast	
EU196717	113	6(X)	Denmark Coast	
EU196718	113	6(X)	Denmark Coast	
Present Study	114	5(F)	Bahia Grande, TX	BG-9
Present Study	115	5(F)	Galveston Bay, TX	BB-7
Present Study	116	5(F)	Galveston Bay, TX	BB-4
Present Study	116	5(F)	Galveston Bay, TX	BB-5
Present Study	116	5(F)	Galveston Bay, TX	BB-6
Present Study	116	5(F)	Galveston Bay, TX	BB-9
Present Study	116	5(F)	Galveston Bay, TX	BB-10
Present Study	116	5(F)	Port Aransas, TX	PA-3
Present Study	116	5(F)	Bahia Grande, TX	BG-8
Present Study	116	5(F)	Port Mansfield, TX	PM-9-1
Present Study	116	5(F)	Port Mansfield, TX	PM-7-1
Present Study	117	5(F)	Port Aransas, TX	PA-2
Present Study	117	5(F)	Port Aransas, TX	PA-5
Present Study	117	5(F)	Port Mansfield, TX	PM-1-1
EU016222	118	5(F)	Atlantic Coast	
EU196730	118	5(F)	Turkey Point, FL	
EU196731	118	5(F)	Turkey Point, FL	
EU196732	118	5(F)	Turkey Point, FL	
DQ431909	119	5(F)	Turkey Point, FL	
EU196725	119	5(F)	Turkey Point, FL	
EU196726	119	5(F)	Turkey Point, FL	
EU196727	119	5(F)	Turkey Point, FL	

EU196728	119	5(F)	Turkey Point, FL	
EU196729	119	5(F)	Turkey Point, FL	
Present Study	120	5(F)	Port Mansfield, TX	PM-11-1
Present Study	121	5(F)	Port Mansfield, TX	PM-8-1
Present Study	122	5(F)	Port Lavaca, TX	PL-3
Present Study	123	5(F)	Port Lavaca, TX	PL-12
Present Study	124	5(F)	Port Lavaca, TX	PL-5
Present Study	125	5(F)	Port Lavaca, TX	PL-6
Present Study	125	5(F)	Galveston Bay, TX	KB-2
Present Study	125	5(F)	Galveston Bay, TX	KB-3
Present Study	125	5(F)	Port Lavaca, TX	PL-4
Present Study	125	5(F)	Port Lavaca, TX	PL-7
Present Study	125	5(F)	Port Lavaca, TX	PL-8
Present Study	125	5(F)	Port Lavaca, TX	PL-10
Present Study	125	5(F)	Port Lavaca, TX	PL-11
Present Study	126	5(F)	Port Lavaca, TX	PL-1
Present Study	127	5(F)	Bahia Grande, TX	BG-7
Present Study	128	5(F)	Bahia Grande, TX	BG-4
DQ431910	129	5(F)	Atlantic Coast	
JF304053	130	5(F)	Atlantic Coast	
JF304058	131	5(F)	Atlantic Coast	
JF304052	132	5(F)	Atlantic Coast	
Present Study	133	5(F)	Port Lavaca, TX	PL-2
JF304054	134	5(F)	Atlantic Coast	
JF304087	136	5(F)	Atlantic Coast	
JF304079	137	5(F)	Atlantic Coast	
JF304081	138	5(F)	Atlantic Coast	
JF304086	139	5(F)	Atlantic Coast	

JF304057	140	5(F)	Atlantic Coast	
JF304031	141	5(F)	Atlantic Coast	
JF304082	142	5(F)	Atlantic Coast	
JF304085	143	5(F)	Atlantic Coast	
JF304056	144	5(F)	Atlantic Coast	
JF304048	145	5(F)	Atlantic Coast	
JF304049	146	5(F)	Atlantic Coast	
JF304059	147	5(F)	Atlantic Coast	
JF304083	148	5(F)	Atlantic Coast	
JF304084	149	5(F)	Atlantic Coast	
JF304078	150	5(F)	Atlantic Coast	
JF304089	151	5(F)	Atlantic Coast	
JF304064	152	5(F)	Atlantic Coast	
JF304046	153	5(F)	Atlantic Coast	
JF304067	154	5(F)	Atlantic Coast	
JF304045	155	5(F)	Atlantic Coast	
JF304076	156	5(F)	Atlantic Coast	
JF304038	157	5(F)	Atlantic Coast	
JF304041	158	5(F)	Atlantic Coast	
JF304039	159	5(F)	Atlantic Coast	
JF304061	160	5(F)	Atlantic Coast	
JF304035	161	5(F)	Atlantic Coast	
JF304040	162	5(F)	Atlantic Coast	
EU274464	163	5(F)	Atlantic Coast	
JF304066	164	5(F)	Atlantic Coast	
JF304036	165	5(F)	Atlantic Coast	
JF304062	166	5(F)	Atlantic Coast	
JF304068	167	5(F)	Atlantic Coast	

JF304063	168	5(F)	Atlantic Coast	
EU274463	169	5(F)	Atlantic Coast	
EU274462	170	5(F)	Atlantic Coast	
EU274456	171	5(F)	Atlantic Coast	
EU274453	172	5(F)	Atlantic Coast	
EU274454	173	5(F)	Atlantic Coast	
EU274459	174	5(F)	Atlantic Coast	
EU274451	175	5(F)	Atlantic Coast	
EU274452	176	5(F)	Atlantic Coast	
EU274444	177	5(F)	Atlantic Coast	
EU274458	178	5(F)	Atlantic Coast	
EU274445	178	5(F)	Atlantic Coast	
EU274447	179	5(F)	Atlantic Coast	
JF304060	180	5(F)	Atlantic Coast	
EU274448	181	5(F)	Atlantic Coast	
JF304091	182	5(F)	Atlantic Coast	
JF304095	183	5(F)	Atlantic Coast	
EU274446	184	5(F)	Atlantic Coast	
EU274455	185	5(F)	Atlantic Coast	
EU274461	186	5(F)	Atlantic Coast	
EU274457	187	5(F)	Atlantic Coast	
EU274449	188	5(F)	Atlantic Coast	
EU274460	189	5(F)	Atlantic Coast	
EU274450	190	5(F)	Atlantic Coast	
JF304042	191	5(F)	Atlantic Coast	
JF304044	192	5(F)	Atlantic Coast	
JF304043	193	5(F)	Atlantic Coast	

BIOGRAPHICAL SKETCH

Nicole Jewel Figueroa was born in Albuquerque, New Mexico. She earned her Bachelor of Science in Fisheries and Wildlife Biology from Oregon State University in 2011. Before attending Oregon State, Nicole was a veterinary technician for eight years and the curator of reptiles at the International Rattlesnake Museum in Albuquerque, New Mexico for two years. This is where she solidified her desire to become a wildlife biologist and pursue a career in science.

After graduation, Nicole dedicated the next three years to starting a family in Tallahassee, Florida while her husband completed his post-doctoral research at Florida State University.

Afterwards, Nicole moved to Brownsville, TX and pursued a master's degree in Biology at the University of Texas Rio Grande Valley. In addition to her Master's thesis, Nicole trained other graduate students in molecular laboratory techniques and bioinformatics, participated in a phylogenetics project on black corals, and assisted in a genome project for multiple marine organisms including several copepod species, coral species, and sea turtles. She earned her Master's in Biology in December of 2017. Nicole plans to dedicate her career to research and education. Nicole can be reached at Nikkijewels84@gmail.com.