Fall 2012

Genetic Population Structure and Cryptic Speciation of Ghost Shrimp (Neotrypaea californiensis) in North American West Coast Estuaries

Daphne Anne Gille
San Jose State University

GENETIC POPULATION STRUCTURE AND CRYPTIC SPECIATION OF GHOST SHRIMP (*NEOTRYPAEA CALIFORNIENSIS*) IN NORTH AMERICAN WEST COAST ESTUARIES

A Thesis

Presented to

The Faculty of the Department of Biological Sciences

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Daphne A. Gille

December 2012

The Designated Thesis Committee Approves the Thesis Titled

GENETIC POPULATION STRUCTURE AND CRYPTIC SPECIATION OF GHOST SHRIMP (*NEOTRYPAEA CALIFORNIENSIS*) IN NORTH AMERICAN WEST COAST ESTUARIES

by

Daphne A. Gille

APPROVED FOR THE DEPARTMENT OF BIOLOGICAL SCIENCES

SAN JOSÉ STATE UNIVERSITY

December 2012

Dr. Leslee A. Parr	Department of Biological Sciences
Dr. J. Brandon White	Department of Biological Sciences
Dr. Joshua A. Mackie	Department of Biological Sciences

ABSTRACT

GENETIC POPULATION STRUCTURE AND CRYPTIC SPECIATION OF GHOST SHRIMP (*NEOTRYPAEA CALIFORNIENSIS*) IN NORTH AMERICAN WEST COAST ESTUARIES

by Daphne A. Gille

Many marine species produce larvae that disperse and develop into post-larvae in the open ocean over a period of weeks to months. However, the patterns and potential of larvae to disperse across long geographic distances are poorly understood. Here genetic variation found at the mitochondrial locus, cytochrome C oxidase subunit I, in adult ghost shrimp, *Neotrypaea californiensis*, was used as a proxy to infer dispersal potential of ghost shrimp larvae found in estuaries along the west coast of the United States. Multiple haplotypes were shared among 346 adult ghost shrimp specimens collected from estuaries in Washington, Oregon, and California indicated that larvae of this species are transported across great distances during the pelagic dispersal phase of development. Interestingly, extreme population structure (F_{ST} ranged from 0.062 to 0.98) was observed among sampling sites located in close proximity to one another. Phylogenetic analyses and analysis of molecular variance revealed that the cause of population structure was in part attributed to the presence of three putative cryptic species that were deeply divergent from *N. californiensis*. Qualitative and quantitative morphological measurements commonly employed for shrimp species identification were inconsistent among putative cryptic species groups but did show that the newly identified cryptic species were not undiagnosed individuals of the closest relative of *N. californiensis*, *N. gigas*.

ACKNOWLEDGEMENTS

I gratefully acknowledge the three members of my thesis committee, Drs. Leslee Parr, Josh Mackie, and Brandon White. I am forever indebted to Dr. Parr who took me on as a graduate student and who has always acted as my advocate, expert advisor, and friend. I thank Dr. Mackie for his input and guidance in the design and execution of this project. I thank Dr. White for years of unwavering encouragement and for offering intellectual, emotional, and even financial support to complete this study when it might have otherwise stalled. I would also like to thank my generous collaborators who collected samples and who were a wealth of knowledge about all things related to burrowing shrimp, Drs. Brett Dumbauld, Ted DeWitt, and Tony D'Andrea. I am also grateful for the laboratory assistance of the undergraduate researchers and volunteers that contributed to this project in the Parr lab, particularly Veronica Chaidez, Cindy Bick, and Mira Brahmbatt.

Above all, I thank the two people whose love, support, and encouragement made the completion of this master's thesis possible: Jennifer Gille (my Mama) and Erin Diaz (my best friend and fiancé!). Thank you for telling me every day that I would get through graduate school and for taking such good care of me. I owe you both several Special Dinners!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Shellfish aquaculture is a burgeoning industry in the United States that significantly contributes to the nation's seafood supply while also boosting the local economy; at present it is estimated that of the shellfish produced globally, 80% is cultured (Shumway 2011). In Washington, Oregon, and California, where estuaries and intertidal mudflats are abundant, the most important and commonly cultivated shellfish is the oyster (Chew 1984). Oyster farming on the west coast of North America was begun in the mid-1800s (Baker 1995) and has been expanded to become a multi-million dollar industry, generating 47% of the country's oysters per year (Dumbauld et al. 2006; U.S. Department of Agriculture 2006). Oyster culture methods require a firm substrate for bedding because without a stable benthos layer, oysters sink into the mud, suffocate, and die (Feldman et al. 2000).

One common problem that threatens the efficiency and success of oyster farming in this region is the presence of the endemic ghost shrimp (*Neotrypaea californiensis*, formerly *Callianassa californiensis*; Manning & Felder 1991), a burrowing decapod thalassinid shrimp species that inhabits estuaries along the eastern Pacific coast ranging from southeast Alaska to Baja California (Stevens 1928; MacGinitie 1934; Dumbauld 1996). *N. californiensis* are strong bioturbators that dig the intricate gallery systems in which they reside (MacGinitie 1934). This behavior causes sediment destabilization, a driving force in benthic community structure (Posey 1986a; 1986b) that also makes ghost shrimp a pest to oyster aquaculture (Dumbauld et al. 2006). Historical attempts to prevent *N. californiensis* settlement in oyster beds by creating a hard bottom substrate

with shell, gravel, buried wooden boards (Stevens 1929), or weighted plastic (Washington Department of Fish & Wildlife 1970) proved largely unsuccessful (Feldman et al. 2000). Current methods used to control *N. californiensis* populations involve the application of insecticides such as carbaryl (Sevin ®) to oyster beds (Dumbauld et al. 2006). Although this approach effectively eliminates ghost shrimp adults and juveniles, it cannot reach larvae that recruit to estuaries via pelagic dispersal (Dumbauld et al. 1996); furthermore, non-specific insecticides also threaten the health of estuary ecosystems. A strategy to manage *N. californiensis* populations at the larval stage without the use of pesticides would therefore be highly desirable. However, such a goal is not easily attained as the dispersal patterns and potential of ghost shrimp larvae are not well understood.

Female burrowing shrimp reproduce by extruding eggs that are then carried on the abdomen for a period of several months, in the case of *N. californiensis*, from March to August (Smith et al. 2008; B. R. Dumbauld, pers. comm.). However, the period of ovigery is believed to vary with both location within an estuary and latitude (Bird 1982; Dumbauld et al. 1996; B. R. Dumbauld, pers. comm.). Once fertilized by a male, eggs are briefly brooded by the female until they are finally released into the estuary where they hatch and develop as larvae during the months of June to August (Dumbauld et al. 1996; B. R. Dumbauld, pers. comm.). It is postulated that larvae then exit the natal estuary on ebb tides and continue development in the nearshore coastal ocean for 6-8 weeks, cycling through five distinct zoeal stages (Fig. 1; McCrow 1972; Johnson & Gonor 1982; Pimentel 1983; unpublished data from A. F. D'Andrea) before recruiting

back into intertidal mudflats as postlarvae by riding nocturnal spring flood tides (Dumbauld et al. 1996; unpublished data from A.F. D'Andrea, B. R. Dumbauld, and K. L. Feldman). Once postlarvae re-enter estuaries, they settle permanently for the remainder of their 4-5 year life span (Bird 1982; Dumbauld et al. 1996).

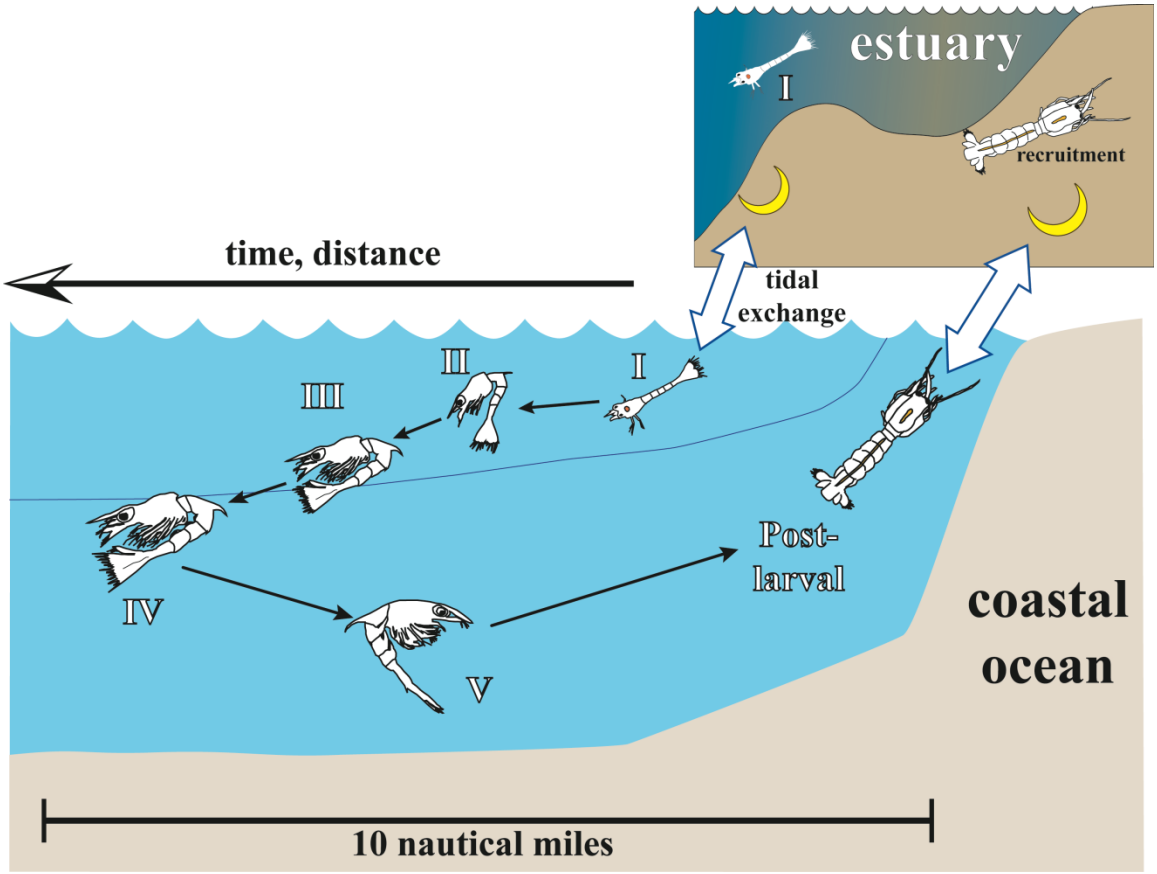


Fig. 1. Five zoeal stages of *N. californiensis* development in relation to position in the water column in the open ocean and approximate distance from the natal estuary (figure courtesy of A.F. D'Andrea). (I) Stage I zoea are 2.8-3.3 mm in length and have a defined rostrum, spine, and medial tooth of telson. (II) Stage II zoea are 3.8-4.4 mm in length and display an additional spine on the medial tooth. (III) Stage III zoea are 4.7-5.2 mm in length with developing uropods. (IV) Stage IV zoea are 5.5-6.3 mm in length and exhibit pleopod buds and fully formed uropods. (IV) Stage V zoea are 6.8-7.5 mm in length and show fully developed pleopods.

There are many natural forces that could act to facilitate the transport of *N. californiensis* larvae great distances offshore once they have entered the open ocean including strong winds and the Coriolis effect (Morgan et al. 2009). These forces act together and cause surface waters (also known as the Ekman layer), where larvae are suspended, to flow perpendicular and away from the coast while also bringing about the upwelling of cold, nutrient-rich waters from the depths (Roughgarden et al. 1988). Marine larvae that disperse in regions of strong upwelling have been shown to be particularly susceptible to offshore transport by this method (Parrish et al. 1981; Yoshioka 1982; Roughgarden et al. 1988; Alexander & Roughgarden 1996). Powerful currents may also significantly influence larval movement in the marine environment. Shanks (2009) provided a comprehensive review of pelagic larval dispersal distances of common benthic marine organisms, demonstrating how currents are capable of carrying larvae several meters to hundreds of kilometers out to sea. It is widely held that the driving force behind long-range larval dispersal along the U.S. Pacific coast excluding Alaska, is the southward flow of the California Current (Shanks et al. 2003; Shanks & Eckert 2005). This current runs > 150 m deep with a broad width of 500 km and slow flow rate of 0.05-0.10 m/s (Strub et al. 1987; Strub & James 2000). Despite the persistent presence of the California Current, it is possible for larvae to recruit to estuaries during periods of upwelling relaxation or reversal (Farrell et al. 1991; Roughgarden et al. 1992).

Long-range transport of *N. californiensis* larvae may also be influenced and even inhibited by larval behavior, pelagic duration, local current variation and coastal topography, and stochasticity (reviewed in Cowen & Sponaugle 2009). While some

larvae rely solely on offshore advection and exhibit passive dispersal behavior larvae may alter their position in the water column by swimming to facilitate transport or retention in response to external cues (Kingsford et al. 2002). Marine larvae can remain nearshore by migrating below the Ekman layer, while ascending to surface waters increases the likelihood of offshore transport (Leis 2006; Leis & McCormick 2002; Morgan et al. 2009). Furthermore, the inherent buoyancy of the larvae may also dictate where it resides in the water column (Cowen 2002). Work by Shanks et al. (2003) showed that larvae disperse at a distance that is significantly correlated with propagule duration, or time spent at sea in the larval stage. Local currents such as retentive eddies, areas that experience reduced flow, or other physical geographic barriers may also prevent larval dispersal over great distances (Sponaugle et al. 2002; Largier 2003; Lipphardt et al. 2006; Sponaugle et al. 2005). For example, Mace and Morgan (2006) found that low-lying headlands inhibit offshore transport of crab larvae along the central California coast.

With such important implications for aquaculture and management, many studies have attempted to decipher the nature and elucidate patterns of *N. californiensis* larval dispersal. Morgan et al. (2009) measured the horizontal distribution of *N. californiensis* larvae offshore of Bodega Bay and Point Reyes, California, the section of the west coast of the U.S. that experiences the strongest upwelling (Koracin et al. 2004; Dorman et al. 2005). *N. californiensis* larvae were found to be most abundant within 3 km from shore and were rarely detected at a distance greater than 6 km from shore despite peak upwelling conditions (Morgan et al. 2009). Morgan and Fisher (2010) examined the distribution of *N. californiensis* throughout the water column 1 km from shore in the

same region. This investigation revealed that *N. californiensis* larvae are most common in bottom waters; however, it is important to note that all samples found at this depth and distance were early stage larvae or postlarvae (Morgan & Fisher 2010). Such data might imply that *N. californiensis* larvae regulate their behavior and swim in such a way as to avoid offshore transport via the shallow Ekman layer.

While distribution surveys point to low dispersal potential of *N. californiensis* larvae, genetic analyses may tell a different story. Because of the difficulty associated with tracking the movements of a large number of larvae, previous studies have employed genetic diversity as a proxy for dispersal distance and potential; it is generally expected that significant genetic population structure will be observed when larval dispersal is restricted among geographic regions (Burton 1983; Bohonak 1999; Kelly & Palumbi 2010). Kozuka (2008) and Buncic (2010) examined mitochondrial DNA (mtDNA) haplotype diversity of *N. californiensis* larvae along the coasts of Washington and Oregon in 2005 and 2006 and found little evidence of genetic differentiation among localities, except among enclosed estuary populations, suggesting that few barriers exist to larval dispersal in these two states. Such results were in line with the expectation that *N. californiensis* would disperse widely given that larvae spend 6-8 weeks developing at sea (Buncic 2010; Dumbauld et al. 1996) and that dispersal distance is highly correlated with pelagic duration (Shanks 2003). Genetic composition also differed between sampling years indicating that signatures of the previous season's larvae were not locally retained thus lending further support to the long-range larval transport hypothesis.

Pernet et al. (2008) tested for the presence of geographic population structure of *N. californiensis* adults and sub-adults collected from sites in Washington, Oregon, and California using two mtDNA loci. The authors determined, much like the larval studies discussed above, that little phylogeographic structure was present in adult populations along the west coast of the U.S (Pernet et al. 2008). Perhaps the most intriguing result of the study was the discovery of a putative cryptic species that appeared as a separate clade ("clade A") in phylogenetic analyses; this group was found only in southern California estuaries and coexisted with *N. californiensis* (Pernet et al. 2008). A second phylogenetic grouping ("clade B") included *N. californiensis* specimens collected from estuaries across the entire sampling range from northern Washington to southern California (Pernet et al. 2008).

The overall objective of the study presented here is to infer *N. californiensis* larval dispersal potential from DNA sequences of adult and sub-adult specimens. A secondary goal is to resolve perceived incongruence between previous survey studies that suggest that *N. californiensis* larvae are often locally retained and genetic analyses that point to the long-range dispersal ability of *N. californiensis* larvae. Specifically, the following question is tested: does significant population structure exist among *N. californiensis* populations throughout Washington, Oregon, and California? This research evaluates mtDNA diversity and expands upon work performed by Pernet et al. (2008) by increasing sampling range and thus population structure resolution and also by including DNA sequences from other closely related shrimp species (*N. gigas*, *N. affinis*, and *Upogebia pugettensis*) in molecular and phylogenetic analyses that have not yet been employed.

Upon the discovery of putative cryptic species here and in Pernet et al. (2008), a second question is addressed: are putative cryptic species merely undiagnosed individuals of an *N. gigas* clade or other closely related shrimp species? In addition to traditional phylogenetic analyses, distinct and reliable morphological differences between *N. californiensis* and *N. gigas* have been identified (Pernet et al. 2010) that make it possible to answer this question.

MATERIALS AND METHODS

Samples. From 2005 to 2007, 346 *N. californiensis* adult samples were opportunistically collected from 32 intertidal estuaries in Washington, Oregon, and California (Table 1). When found, samples of other thalassinid shrimp species including *N. gigas* (giant ghost shrimp, distinguished at the time of collection by its distinctive large size), *N. affinis* (tidepool ghost shrimp), and *Upogebia pugettensis* (blue mud shrimp) were also collected (Table 2). The species, gender, carapace length (to the nearest mm) of each individual was noted at the time of collection. Shovels and yabby pumps were used to harvest all shrimp samples. Shrimp were immediately frozen at -20°C upon return to the laboratory and were stored in freezers until thawing for DNA extraction.

Table 1. *N. californiensis* sample collection locations, quantities, and dates.

	Location	ID	n	Date of collection	Latitude N	Longitude W
Washington	Pysht River Estuary	PRE	5	6/4/07	48.20	124.10
	Budd Inlet	BI	6	5/7/07-5/15/07	47.07	122.90
	Gray's Harbor	GH	12	9/10/06	46.98	124.10
	Potlatch State Park	PSP	9	11/1/06	47.35	123.16
	Ellen Sands	ES	11	7/11/05	46.66	123.99
	Smoky Hollow	SH	11	8/20/05	46.42	123.99
	Stackpole	ST	11	8/21/05	46.62	124.04
	Firecracker Point	FP	11	9/19/05	46.90	124.09
	False Bay	FB	11	2/16/06	48.49	123.07
	Puget Sound	PS	5	7/29/06	47.10	122.73
	Columbia River	CR	11	6/19/05-8/30/05	46.17	124.25
Oregon	Yaquina Bay Seawall	SW	11	8/6/05	44.63	124.05
	Yaquina Bay	YB	11	8/3/05	44.59	123.96
	Yaquina Bay Island	YI	11	8/5/05	44.58	123.96
	Umpqua	UQ	11	6/23/05	43.72	124.15
	Tillamook	TM	11	7/20/05-7/21/05	45.56	123.92
	Coos Bay	CB	11	8/5/05-8/9/05	43.36	124.30
	Alsea Bay	AB	11	8/11/06	44.43	124.07
	Nestucca	NS	11	9/1/06	45.18	123.95
	Cape Meares	CM	11	6/20/05-8/31/05	45.50	124.12
	Cascade Head	CA	11	6/21/05	45.05	124.20
	Newport Harbor	NH	10	8/29/05	44.62	124.25
	Heceta Head	HH	10	8/21/05	44.13	124.20
California	Crescent City	Cres	7	4/19/07	41.75	124.19
	West Bodega Bay	WBH	4	7/29/06	38.32	123.05
	East Bodega Bay	EBH	5	7/29/06	38.33	123.04
	Morro Bay	MR	1	9/9/06	35.35	120.84
	Simple Green Mudflat	SG	8	9/9/06	44.62	124.04
	San Pedro	SP	18	9/15/06	33.71	118.28
	Catalina Harbor	CH	12	11/3/06	33.44	118.50
	San Diego	SD	33	4/30/06	32.76	117.24
	Mission Bay	MB	25	4/30/06	32.76	117.25
	Total		**346**			

Table 2. *N. gigas*, *N. affinis*, and *U. pugettensis* sample collection locations, quantities, and dates.

	Location	Species	ID	n	Date of collection	Latitude N	Longitude W
WA	Pysht River Estuary	U. pugettensis	PRE_Upu	1	6/4/07	48.20	124.10
	Gray's Harbor	U. pugettensis	GH_Upu	2	9/10/06	46.98	124.10
	Willapa Bay	U. pugettensis	WBG_Upu	2	8/21/05	46.62	124.04
	Gray's Harbor	N. gigas	GH_Ngi	2	9/10/06	46.98	124.10
	Ellen Sands	N. gigas	ES_Ngi	1	7/11/05	46.66	123.99
OR	Yaquina Bay	N. gigas	YB_Ngi	3	8/10/06	44.62	124.04
CA	Santa Barbara	N. affinis	SBB_Naf	2	10/5/06	34.41	119.88
	Catalina Harbor	U. pugettensis	CH_Upu	14	11/3/06	33.43	118.51

DNA Extraction. Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) according to the manufacturer's instructions with modifications. Briefly, a small section of the uropod (10-30 μg) was excised with a scalpel and submerged in a mixture of 500 μl of Nuclei Lysis Solution and 120 μl of 0.5M EDTA. The tissue was homogenized with a small pestle and chilled on ice. An additional volume of 600 μl of the Nuclei Lysis Solution and EDTA mixture was applied to the sample as well as 17.5 μl of proteinase K (20 mg/ml; Fisher Scientific, Pittsburgh, PA). Samples were incubated overnight at 55°C with gentle mixing. Proteins were then removed following the addition of 200 μl of Protein Precipitation Solution and high-speed centrifugation. DNA was isolated and washed with 600 μl of 100% isopropanol and 600 μl of 70% ethanol. Extracted DNA was reconstituted in 100 μl of DNA Rehydration Solution and stored at -20°C until it was thawed for use in amplification reactions.

DNA Amplification. Polymerase chain reactions (PCR) were performed to amplify a 520 bp region located in the middle of the mitochondrial gene encoding cytochrome C oxidase subunit I (COI). Each 25 μl amplification reaction contained a buffered solution of 0.2 mM dNTPs, 2mM $MgCl_2$, 0.3 mg/ml bovine serum albumin, 0.2 μM forward primer SCOI-F-NEW (5'-CCTGGGTTTGGTATAATTTCTCA-3'), 0.2 μM forward primer SCOI-R-NEW (5'-ATCGGGGTAATCTGAATATCG-3'), 0.5 U Taq polymerase, and 10-30 ng of template DNA. A negative control consisting of water rather than template DNA was included in each PCR run. Thermal cycler conditions for the amplification of the COI segment were: 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 58-62°C for 30 s, and 72°C for 1 min, and a final extension of 72°C for 10 min.

DNA Sequencing. PCR products that showed a discrete band on 2% agarose gel, indicating that the appropriate 520 bp segment had been amplified, were purified with QIAquick spin columns (Qiagen, Chatsworth, CA); remaining primer sequence, dNTPs, and incomplete amplification products were removed by this method. Purified PCR products were sequenced in the forward direction using primer SCOI-F-NEW at Geneway Research (Hayward, CA). Sequences were obtained using the ABI Prism 3700 DNA Analyzer and automated sequencer (Applied Biosystems, Foster City, CA) using BigDye™ terminator methodology. Sequences containing ambiguities were sequenced in the reverse direction using primer SCOI-R-NEW.

Molecular Data Analysis. Sequences were edited using Sequencher v. 4.7 (Gene Codes Corporation, Ann Arbor, MI) software and multiple sequence alignment was executed on a 393-520 bp COI segment using Clustal X v. 2.1 (Larkin 2007). TCS v. 1.21 (Clement et al. 2000) was used to identify redundant haplotypes and to calculate haplotype frequency and distribution for samples of all four species. DnaSP v. 5.10.01 (Librado & Rozas 2009) was used to estimate molecular diversity parameters including haplotype (*He*) and nucleotide (π) and diversity and the average number of nucleotide differences at each sampling site (*k*) of all *N. californiensis* sequences.

Tajima's *D*, Fu's F_S, and respective significance values were calculated for each population of *N. californiensis* using Arlequin v. 3.11 (Excoffier et al. 2005); samples from SW, YB, and YI were grouped together because of the close proximity of the sampling locations. These indices are used to test the null hypothesis that patterns of *N. californiensis* COI sequence variation are selectively neutral. Tajima's *D* is estimated from pairwise comparisons of the number of observed polymorphic nucleotide sites to the average number of nucleotide differences among all sampled sequences (Tajima 1989; Ford 2002). Fu's F_S estimates the probability of obtaining the observed number of alleles or more in a given sample size of sequences given the observed level of pairwise differences under the infinite sites model (Fu 1997; Excoffier et al. 2005). The significance of both Tajima's *D* and Fu's F_S was obtained in Arlequin by comparing 1000 runs of a coalescent simulation algorithm adapted by Hudson (1990) under the assumption of population equilibrium and selective neutrality to the observed distribution (significant when $p < 0.05$; Excoffier et al. 2005). It is necessary to perform both

statistical tests because a single index alone likely has inadequate power to capture all possible evolutionary processes (Ferretti et al. 2010).

Arlequin v. 3.11 (Excoffier et al. 2005) was also employed to determine the extent of *N. californiensis* population subdivision. Population pairwise comparison F_{ST} values and a genetic matrix were computed for all pairs of *N. californiensis* populations and significance was assessed after 1000 permutations. F_{ST} is a measure of genetic differentiation that is used to infer population structure and is estimated by comparing the haplotype diversity of a subpopulation to that of the entire population (Wright 1951; 1965). An F_{ST} value of 0 implies the subpopulations have the same allele frequencies and that a subsample population is indistinguishable from a selection of variation from the entire population (Meirmans & Hedrick 2011). Conversely, values tending toward 1 indicate the presence of different unique fixed alleles and distinct subpopulations (Meirmans & Hedrick 2011). If F_{ST} is statistically significant ($p < 0.05$) and greater than zero, the null hypothesis of no population structure between groups is rejected. The discovery of three highly divergent clades that potentially represented cryptic species in addition to a large *N. californiensis* clade (see Results) prompted the estimation of pairwise F_{ST} a second time with samples categorized by location and also by putative cryptic species grouping (see *Morphological Analysis*, Table 3).

Pairwise distances among sampling sites in meters were calculated from decimal latitude and longitude values and used to create a physical distance matrix by The Geographic Distance Matrix Generator v. 1.2.3 (Ersts 2012). Finally, a Mantel test (Mantel 1967) that involves the plotting of genetic distance among population pairs

against geographic distance was performed in Arlequin with 1,000 permutations to determine whether localities are genetically distinct, also known as isolation by distance (Wright 1943). It is expected that populations located in close proximity to one another will show greater genetic similarity than geographically distant populations. The output of a Mantel test is the correlation coefficient of the two matrices (*r*) that indicates that strength of the isolation by distance relationship.

Additionally, an analysis of molecular variance (AMOVA) was performed to evaluate variance partitioning within and among each of the three putative cryptic species and the large *N. californiensis* group (see *Morphological Analysis*, Table 3). This hierarchical AMOVA was conducted in Arlequin using standard non-parametric permutation procedures (haplotypic format with 1000 permutations; Excoffier et al. 1992). The null hypothesis is that no significant genetic breaks exist among the three cryptic species groups and the large group of *N. californiensis* specimens.

Phylogenetic Analysis. Multiple methods were utilized to infer phylogenetic relationships among *N. californiensis*, *N. gigas*, *N, affinis*, and *U. pugettensis* samples. Maximum likelihood rapid bootstrapping analyses were conducted on non-redundant *N. californiensis*, *N. gigas*, *N, affinis*, and *U. pugettensis* sequences using the default parameters in the program Randomized Axelerated Maximum Likelihood (RAxML; Stamatakis 2006; Stamatakis et al. 2008) v. 7.2.8 via the Cyberinfrastructure for Phylogenetic Research (CIPRES) portal ("XSEDE" version; Miller et al. 2010) with 100 bootstrap replicates. A GTR + CAT model was employed to calculate bootstrap values and a GTR + GAMMA model of nucleotide substitution was used for final tree inference.

The following samples were selected as the collective outgroup: CH4_Upu, CH9_Upu, CH13_Upu, CH14_Upu, CH15_Upu, CH22_Upu, Hap20_Upu, Hap40_Upu, Hap21_Upu. Phylogenetic tree construction using the maximum likelihood method involves the selection of the tree within the search space that has the highest score based upon the optimization of branch lengths (Felsenstein 1981; Yang and Rannala 2012). Maximum likelihood phylogenetic tree construction was selected for use in this study as a means to corroborate findings inferred from haplotype networks and because such an approach is generally the most robust and accurate method of predicting a true evolutionary tree (Tateno et al. 1994; Huelsenbeck 1995; Kuhner and Felsenstein 1995; Huelsenbeck and Rannala 1997).

The construction of a 95% statistical parsimony network of haplotypes was performed in TCS v. 1.21 with gaps set as missing characters (Templeton et al. 1992; Clement et al. 2000). Alignment files produced by Clustal X were modified for use in TCS using ALTER (Glez-Peña et al. 2010). The TCS algorithm calculates the maximum number of differences required to establish a parsimonious connection between two haplotypes with 95% probability (i.e. the parsimony connection limit; Clement et al. 2000). Haplotypes with a single difference between them are joined; this process is repeated for two differences and up until a single network connecting all haplotypes has been constructed or until the 95% parsimony cut-off has been reached (Templeton et al. 1992; Clement et al. 2000; Posada and Crandall 2001). In the graphical output, redundant sequences are collapsed into a single haplotype and the size of each distinct node corresponds to the frequency of the haplotype. The most likely ancestral haplotypes

are displayed in a rectangle rather than an oval. Haplotype networks allow for and enable the visualization of reticulate evolution and non-bifurcating relationships and are therefore often preferred to other traditional methods of phylogenetic reconstruction (Posada and Crandall 2001).

Morphological Analysis. Both molecular data and phylogenetic analyses revealed the presence of multiple putative cryptic species (see Results). Putative cryptic species groups were so named if the sequences grouped the specimens together in a distinct phylogenetic clade and also formed a unique haplotype network (Table 3); three of these groupings were identified. If a population included individuals thought to belong to one of three putative cryptic species, all shrimp samples from the location were considered for morphological analysis. Individuals belonging to one putative cryptic species grouping, Group 1, were found in SG, SP, and CH (Table 3). All samples from BI, PRE, and GH_Ngi grouped together as a second putative cryptic species group, Group 2, and several samples from CA and CR made up a third putative cryptic species group, Group 3 (Table 3). In numerous instances, shrimp samples were too degraded for analysis or were no longer in existence and were therefore not included in the morphological analysis (Table 3).

Table 3. Three putative cryptic species groups based upon the results of haplotype network and phylogenetic analyses. Samples thought to belong to a cryptic species are listed first; other samples thought to be *N. californiensis* found at the same locations are listed second. Morphological measurements were performed on samples highlighted in bold. Samples not highlighted in bold were too degraded for analysis or no longer existed in the laboratory.

Group 1:	**SG20**, **SG21**, **SG22**, SG23, **SG24**, **SG25**, SG26, **SG28**, SP3, **SP6**, **SP13**, SP14, SP15, SP17, SP19, SP22, **SP24**, **CH2**, **CH3**, **CH4**, **CH6**, **CH8**, **CH9**, CH10, CH11, CH12, CH13, CH14, CH17
N. californiensis Group 1:	**SP1**, **SP2**, **SP4**, **SP5**, **SP8**, SP10, SP11, **SP12b**, SP18
Group 2:	PRE1_1, PRE1_3, PRE1_4, PRE2_3, PRE2_7, **BI1_4**, BI3_3, **BI3_5**, BI3_9, BI4_3, BI4_5,GH206_Ngi, GH207_Ngi
N. californiensis Group 2:	None
Group 3:	CR4066, CR4067, CR7002, CR7003, CR7006, CA1026, CA1029
N. californiensis Group 3:	CR4064, CR4065, CR4068, CR7001, CR7004, CR7005, CA1019, CA1020, CA1021, CA1022, CA1023, CA1024, CA1025, CA1027, CA1028

To confirm that these specimens were not merely undiagnosed groupings of *N. gigas*, a close relative that is similar in appearance to *N. californiensis*, qualitative and quantitative measurements known to be useful in distinguishing *N. californiensis* and *N. gigas* (Pernet et al. 2010) were collected. The measurements of eyestalk length and shape have proven to be the most robust means currently available to differentiate the two species that can be employed for both sexes and all sizes of shrimps (Pernet et al. 2010). The eyestalks of *N. californiensis* have a convex outer edge and extend to the base or up to 1/4 the length of the second segment of the first antennae, whereas *N. gigas* eyestalks have a concave outer edge and extend 1/3 to 3/4 the length of the antennae (Pernet et al. 2010). The ratio of the length and width of the carpus of the male claw may also be used to separate the two species, the ratio being generally larger in *N. californiensis* (Hart 1982; Pernet et al. 2010). However, this trait is less robust and decidedly less reliable

because it may only be used to discriminate males of the two species with a carapace greater than 10 mm in length and because the major claw is often lost during sample harvest.

Eyestalk length was examined under 10x microscope magnification and recorded as extending to the base, 1/4, 1/3, or 1/2 the length of the second segment of the first antenna for all samples in bold (Table 3) except for one whose eyestalks were damaged upon collection (SG28); eyestalk shape was simultaneously scored as concave, convex, or unknown. When present, the width and length of the carpus of the male claw were measured using calipers to the nearest millimeter. Carapace length, recorded by collectors at the time of sample harvest using calipers to the nearest tenth of a millimeter, was also included in morphological analyses. The ratio of the carpus width and length was plotted against carapace length and displayed in a scatter plot to distinguish *N. californiensis* from *N. gigas* according to the methods described by Pernet et al. (2010).

RESULTS

DNA sequencing of a 520 bp region of the cytochrome C oxidase subunit I (COI) gene of 373 shrimp samples collected from the eastern Pacific coast of the United States showed 154 polymorphic sites that yielded 169 unique haplotypes, 38 of which were shared by two or more individuals (Table 4). Mean nucleotide base frequencies across all sites were A: 23.42%, C: 14.94%, G: 21.00%, and T: 40.64%. The three most common haplotypes (Hap4, Hap5, and Hap8) were found throughout the sampling range in Washington, Oregon, and California as was one other (Hap10; Table 4). Washington and Oregon shared three additional haplotypes (Hap6, Hap13, and Hap14) between them,

while Washington and California shared one other haplotype (Hap9; Table 4). No haplotypes were found in both states of Washington and Oregon only. Two haplotypes (Hap2 and Hap4) appeared in both *N. californiensis* and samples identified as *N. gigas* at the time of collection. One haplotype (Hap15) belonged to only *N. gigas* samples and another (Hap22) to only *U. pugettensis* samples; interestingly, all *U. pugettensis* individuals displayed the same haplotype. All other haplotypes were singletons or specific to the sampling location or state (Table 4).

The average number of nucleotide differences between haplotypes (k) across all locations was 8.245; this value among individual sites ranged greatly from 1.533 (BI) to 30.431 (SP; Table 5). Overall haplotype diversity (h) was high at 0.877 while nucleotide diversity (π) was low at 0.017; the sets of values for each were similar across locations (Table 5).

The results of neutrality tests using all *N. californiensis* samples were incongruent. Tajima's D values for individual populations were mixed positive and negative and generally not statistically significant ($p > 0.05$) with the one exception of FP ($p = 0.037$; Table 6). Fu's F_S was largely negative for all populations excepting Cres ($F_S = 1.992$) and EBH ($F_S = 0.127$), neither of which were significant ($p = 0.557$ and $p = 0.327$, respectively); the majority of the remaining F_S were significant (Table 6).

Table 4. Distribution and frequency of *N. californiensis* haplotypes found in two or more individuals in more than one sampling location. Starred haplotypes indicate those found in both *N. californiensis* and *N. gigas* samples.

		Haplotypes																																					
	Location	1	2*	3	4*	5	6	7	8	9	10	11	12	13	14	16	17	18	19	20	21	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Washington	PRE	2																																			
	BI	2	1																																		
	GH			1	2	5	1	2																													
	PSP				2				2																												
	ES				3	3			1	1	1																										
	SH				3	6			2																												
	ST				4	4			3																												
	FP			1	1	3						2																									
	FB				2	6			1																												
	PS					1			1				2																								
	CR													1	3																						
Oregon	SW				4	2			1																												
	YB				4	3																															
	YI					3	1				1					1																						
	UQ				3	2			4																												
	TM				3	4			1																												
	CB				3	4			1							2	1																				
	AB				1	4			4																												
	NS				4	4											1																				
	CM				2	1								1																							
	CA					1									1																						
	NH				3																																
	HH				4				1																												
California	Cres				1																																
	WBH																																				
	EBH									1								1																			
	MR										1																										
	SG																		4	1																	
	SP								2									1	5		2																
	CH																		5	2																	
	SD				1				4													2	2	2	2	2	2	2	2	2						
	MB					1			4																					1	2	2	2	2	2	2	2
	TOTAL	4	1	2	50	57	2	2	32	2	3	2	2	2	4	3	2	2	14	3	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2

Table 5. Number of samples collected per location and the resulting number of haplotypes, haplotype diversity (h), nucleotide diversity (π), and average number of nucleotide differences between haplotypes found at each.

	Site ID	Number of sequences	Number of haplotypes	h	π	k
Washington	PRE	5	4	0.900	0.006	2.800
	BI	6	4	0.800	0.003	1.533
	GH	12	6	0.818	0.013	6.848
	PSP	9	7	0.944	0.024	11.222
	ES	11	7	0.891	0.009	4.582
	SH	11	3	0.655	0.009	4.473
	ST	11	3	0.727	0.009	4.509
	FP	11	8	0.927	0.015	7.636
	FB	11	5	0.709	0.010	5.200
	PS	5	4	0.900	0.008	4.200
	CR	11	9	0.945	0.040	20.873
Oregon	SW	11	7	0.873	0.012	5.982
	YB	11	6	0.836	0.009	4.764
	YI	11	9	0.945	0.009	4.545
	UQ	11	5	0.818	0.009	4.836
	TM	11	6	0.836	0.009	4.909
	CB	11	5	0.818	0.008	4.218
	AB	11	5	0.782	0.009	4.509
	NS	11	5	0.782	0.009	4.545
	CM	11	10	0.982	0.035	18.073
	CA	11	11	1.000	0.044	23.091
	NH	10	8	0.933	0.018	9.444
	HH	10	7	0.867	0.021	10.978
California	Cres	7	7	1.000	0.026	10.286
	WBH	4	4	1.000	0.009	4.333
	EBH	5	5	1.000	0.033	14.400
	MR	1	1	N/A	N/A	N/A
	SG	8	5	0.786	0.009	4.357
	SP	18	12	0.922	0.069	30.431
	CH	12	7	0.833	0.011	5.258
	SD	33	22	0.977	0.012	6.125
	MB	25	16	0.970	0.013	6.647
	Mean			0.877	0.017	8.245

Table 6. Results of Tajima's *D* and Fu's F_S neutrality tests and associated p-values for individual populations of *N. californiensis*. Starred values indicate significance ($p < 0.05$).

	Site ID	Tajima's *D*	Tajima's *D* p-value	Fu's F_S	Fu's F_S p-value
Washington	PRE	-1.162	0.074	-2.371	0.027*
	BI	-0.676	0.338	-5.148	0.000*
	GH	-0.913	0.199	-6.567	0.001*
	PSP	-0.081	0.505	-2.535	0.057
	ES	0.141	0.601	-7.447	0.000*
	SH	1.915	0.992	-7.569	0.000*
	ST	1.964	0.983	-7.528	0.000*
	FP	-1.594	0.037*	-5.169	0.005*
	FB	0.068	0.570	-6.829	0.001*
	PS	0.661	0.711	-1.633	0.092
	CR	0.939	0.851	-2.171	0.085
Oregon	SW	0.424	0.696	-6.186	0.000*
	YB	1.155	0.884	-7.253	0.000*
	YI	0.905	0.826	-7.487	0.000*
	UQ	0.784	0.786	-7.178	0.001*
	TM	0.861	0.841	-7.106	0.002*
	CB	1.567	0.953	-7.870	0.000*
	AB	0.863	0.829	-7.528	0.000*
	NS	1.411	0.921	-7.487	0.000*
	CM	1.366	0.945	-2.503	0.071
	CA	0.884	0.867	-1.953	0.117
	NH	-0.664	0.246	-3.650	0.026*
	HH	-0.886	0.220	-3.218	0.037*
California	Cres	0.120	0.500	1.992	0.557
	WBH	-0.069	0.629	-0.715	0.162
	EBH	1.489	0.968	0.127	0.327
	MR	N/A	N/A	N/A	N/A
	SG	-0.295	0.407	-4.335	0.003*
	SP	1.891	0.988	-4.257	0.033*
	CH	0.251	0.634	-7.906	0.000*
	SD	-0.883	0.193	-25.259	0.000*
	MB	-0.901	0.171	-23.414	0.000*

A wide range of significant pairwise F_{ST} values from 0.077 (GH and SD) to 0.956 (BI and PS; Table 7A) was observed among all *N. californiensis* populations. Of these, 71% showed very great genetic differentiation ($F_{ST} > 0.25$), 18% were greatly differentiated ($0.25 > F_{ST} \geq 0.15$), and 11% were moderately differentiated ($0.15 > F_{ST} \geq 0.05$). The least amount of population structure was found among sites in Washington with the exception of two, PRE and BI that had high pairwise F_{ST} values when compared to all other locations (Table 7A). Large pairwise F_{ST} values were also detected between Cres, WBH, and EBH and all other populations, CR and all others, as well as SG, SP, and CH and all others (Table 8A). A Mantel test indicated a weak positive ($r = 0.132$) but non-significant ($p = 0.13$) relationship between genetic differentiation and geographical distance. When divided by location and by putative cryptic clade, a similar wide range of significant pairwise F_{ST} values from 0.062 (UQ and SD) and 0.976 (CR cryptic Group 3 and BI cryptic Group 2) was obtained (Table 7B). CA and SP sites included individuals that belonged to cryptic species clades as well as the large overall *N. californiensis* group. When divided, smaller pairwise F_{ST} values were generally found between the *N. californiensis* samples and all other locations whereas cryptic species samples remained high (Table 7B).

Table 7. Pairwise F_{ST}. (A) All *N. californiensis* populations. (B) All *N. californiensis* populations with samples divided according to cryptic species clade (c1 = clade 1, c2 = clade 2, and c3 = clade 3 described in Table 3). Starred values are significant ($p < 0.05$); bold values are highly significant ($p < 0.001$). Table 7A continued on next page.

(A)

	PRE	BI	GH	PSP	ES	SH	ST	FP	FB	PS	CR	SW	YB	YI	UQ	TM
PRE	0.000															
BI	0.051	0.000														
GH	**0.915**	**0.925**	0.000													
PSP	**0.864**	**0.880**	0.051	0.000												
ES	**0.936**	**0.945**	0.097*	0.058	0.000											
SH	**0.937**	**0.946**	0.003	0.065	0.025	0.000										
ST	**0.936**	**0.945**	0.058	0.037	-0.068	-0.043	0.000									
FP	**0.910**	**0.921**	-0.039	0.082	0.168*	0.026	0.118	0.000								
FB	**0.933**	**0.942**	0.006	0.098	0.081	-0.075	0.017	0.026	0.000							
PS	0.942	**0.956**	0.003	0.056	0.202*	-0.018	0.108	-0.030	-0.022	0.000						
CR	**0.733**	**0.752**	**0.499**	**0.374**	**0.538**	**0.540**	**0.538**	**0.489**	**0.539**	**0.459**	0.000					
SW	**0.921**	**0.931**	0.015	0.052	-0.013	-0.076	-0.060	0.048	-0.049	0.014	**0.526**	0.000				
YB	**0.936**	**0.945**	0.054	0.068	-0.040	-0.016	-0.050	0.106	0.026	0.155	**0.543**	-0.039	0.000			
YI	0.936*	0.944*	-0.004	0.105*	0.121*	0.008	0.076	0.020	0.011	0.029	**0.534**	0.028	0.084	0.000		
UQ	**0.932**	**0.942**	0.056	0.015	-0.049	-0.018	-0.083	0.117	0.048	0.097	**0.526**	-0.032	-0.018	0.083	0.000	
TM	**0.934**	**0.943**	0.015	0.090	0.087	-0.083	0.011	0.013	-0.077	-0.042	**0.538**	-0.052	0.026	0.023	0.037	0.000
CB	**0.940**	**0.948**	0.021	0.050	-0.052	-0.052	-0.080	0.085	-0.003	0.099	**0.543**	-0.057	-0.054	0.032	-0.060	0.001
AB	**0.936**	**0.945**	0.019	0.053	0.074	-0.070	-0.009	0.037	-0.039	-0.059	**0.534**	-0.037	0.063	0.025	-0.008	-0.054
NS	**0.936**	**0.945**	0.050	0.064	-0.052	-0.028	-0.070	0.111	0.016	0.142	**0.543**	-0.056	-0.079	0.088	-0.036	0.018
CM	**0.807**	**0.824**	0.114	0.030	0.188*	**0.183**	**0.181**	0.131*	0.181	0.137	**0.295**	**0.168**	0.189*	0.184*	**0.174**	**0.187**
CA	**0.742**	**0.762**	**0.269**	0.168	**0.332**	**0.322**	**0.326**	**0.265**	0.322	0.227	0.059	**0.314**	**0.337**	**0.313**	**0.311**	**0.319**
NH	**0.898**	**0.911**	0.068	0.017	0.039	0.129*	0.068	0.125	0.160*	0.200*	**0.450**	0.088	0.039	0.155*	0.065	0.167*
HH	**0.885**	**0.899**	0.149	-0.016	0.038	**0.148**	0.049	**0.209**	0.197*	0.198*	0.432*	0.103*	0.093*	**0.205**	0.029	0.193*
Cres	**0.289**	**0.324**	**0.231**	**0.154**	**0.226**	**0.227**	**0.225**	**0.216**	**0.228**	0.089	**0.209**	**0.220**	**0.225**	**0.228**	**0.223**	**0.228**
WBH	0.928*	0.947*	**0.767**	0.588*	**0.830**	**0.830**	**0.829**	**0.756**	**0.822**	0.832*	**0.345**	**0.795**	**0.830**	**0.825**	**0.817**	**0.824**
EBH	0.829*	0.856*	0.372*	0.115	0.446*	0.461*	0.447*	**0.375**	**0.464**	0.380*	0.207*	**0.424**	**0.458**	**0.456**	0.422*	**0.465**
MR	0.948	0.972	-0.107	-0.578	-0.527	-0.004	-0.305	0.026	0.067	0.300	0.207	-0.242	-0.260	0.038	-0.364	0.088
SG	0.894*	0.913*	**0.914**	**0.871**	**0.930**	**0.931**	**0.930**	**0.909**	**0.928**	**0.932**	**0.757**	**0.918**	**0.930**	**0.930**	**0.927**	**0.929**
SP	**0.509**	**0.536**	**0.490**	**0.397**	**0.509**	**0.507**	**0.506**	**0.479**	**0.508**	**0.434**	**0.294**	**0.502**	**0.512**	**0.505**	**0.498**	**0.507**
CH	**0.875**	**0.891**	**0.909**	**0.873**	**0.922**	**0.923**	**0.923**	**0.905**	**0.921**	**0.921**	**0.773**	**0.913**	**0.923**	**0.922**	**0.920**	**0.922**
SD	**0.919**	**0.924**	0.077*	**0.087**	**0.113**	0.137*	**0.106**	**0.133**	**0.182**	0.163*	**0.597**	**0.135**	**0.130**	0.124*	0.062	**0.179**
MB	**0.911**	**0.917**	0.079*	0.049	0.080*	0.107*	0.061	**0.143**	0.160*	0.123*	**0.551**	0.103*	0.121*	0.117*	0.014	0.150*

(A) Table 7A continued from previous page.

	CB	AB	NS	CM	CA	NH	HH	Cres	WBH	EBH	MR	SG	SP	CH	SD	MB
CB	0.000															
AB	-0.008	0.000														
NS	-0.063	0.041	0.000													
CM	**0.179**	**0.184**	**0.182**	0.000												
CA	**0.326**	**0.311**	**0.336**	0.109	0.000											
NH	0.059	**0.172**	0.047	0.074	**0.244**	0.000										
HH	0.086	0.156*	0.070	0.111*	0.248*	0.023	0.000									
Cres	**0.226**	**0.226**	**0.225**	**0.178**	**0.182**	0.188*	**0.183**	0.000								
WBH	**0.837**	**0.826**	**0.831**	0.439*	**0.378**	**0.716**	**0.676**	0.070	0.000							
EBH	**0.453**	**0.450**	**0.458**	0.042	0.100	**0.276**	0.229	0.035	0.197	0.000						
MR	-0.190	0.064	-0.250	-0.357	-0.169	-0.484	-0.748	-0.772	0.816	-0.194	0.000					
SG	**0.933**	**0.930**	**0.931**	**0.821**	**0.762**	**0.900**	**0.888**	**0.385**	**0.916**	**0.843**	0.928	0.000				
SP	**0.508**	**0.500**	**0.512**	**0.369**	**0.304**	**0.457**	**0.439**	**0.253**	0.263*	0.235	0.203	0.305*	0.000			
CH	**0.925**	**0.922**	**0.923**	**0.830**	**0.778**	**0.897**	**0.888**	**0.454**	**0.902**	**0.849**	0.912	-0.067	**0.338**	0.000		
SD	0.087*	**0.128**	0.127*	**0.243**	**0.378**	0.098*	0.129*	**0.429**	**0.785**	**0.442**	-0.144	**0.917**	**0.594**	**0.914**	0.000	
MB	0.063	0.085*	0.105*	**0.196**	**0.327**	0.105*	0.059	**0.370**	**0.762**	**0.397**	-0.289	**0.909**	**0.551**	**0.906**	0.017	0.000

(B) Table 7B continued on next page.

	PRE _c2	BI_c2	GH	PSP	ES	SH	ST	FP	FB	PS	CR	CR_c3	SW	YB	YI	UQ	TM
PRE _c2	0.000																
BI_c2	0.025	0.000															
GH	**0.915**	**0.925**	0.000														
PSP	**0.864**	**0.880**	0.051	0.000													
ES	**0.936**	**0.945**	0.097	0.058	0.000												
SH	**0.937**	**0.945**	0.003	0.065	0.025	0.000											
ST	**0.936**	**0.945**	0.058	0.037	-0.068	-0.043	0.000										
FP	**0.910**	**0.921**	-0.039	0.082	0.168*	0.026	0.118	0.000									
FB	**0.933**	**0.942**	0.006	0.098	0.081	-0.075	0.017	0.026	0.000								
PS	0.942*	0.956*	0.003	0.056	0.202	-0.018	0.108	-0.030	-0.022	0.000							
CR	0.846*	**0.870**	**0.527**	0.319	**0.604**	**0.599**	**0.598**	**0.516**	**0.594**	0.520*	0.000						
CR_c3	0.963*	**0.976**	**0.866**	**0.780**	**0.905**	**0.908**	**0.907**	**0.859**	**0.901**	0.933*	0.742*	0.000					
SW	**0.921**	**0.931**	0.015	0.052	-0.013	-0.076	-0.060	0.048	-0.049	0.014	**0.564**	**0.881**	0.000				
YB	**0.936**	**0.945**	0.054	0.068	-0.040	-0.016	-0.050	0.106	0.026	0.155	**0.606**	**0.907**	-0.039	0.000			
YI	**0.936**	**0.944**	-0.004	0.105*	0.121	0.008	0.076	0.020	0.011	0.029	**0.592**	**0.905**	0.028	0.084	0.000		
UQ	**0.932**	**0.941**	0.056	0.015	-0.049	-0.018	-0.083	0.117*	0.048	0.097	**0.578**	**0.899**	-0.032	-0.018	0.083	0.000	
TM	**0.934**	**0.943**	0.015	0.090	0.087	-0.083	0.011	0.013	-0.077	-0.042	**0.593**	**0.903**	-0.052	0.026	0.023	0.037	0.000
CB	**0.940**	**0.948**	0.021	0.050	-0.052	-0.052	-0.080	0.085	-0.003	0.099	**0.608**	**0.913**	-0.057	-0.054	0.032	-0.060	0.001
AB	**0.936**	**0.944**	0.019	0.053	0.074	-0.070	-0.009	0.037	-0.039	-0.059	**0.590**	**0.905**	-0.037	0.063	0.025	-0.008	-0.054
NS	**0.936**	**0.945**	0.050	0.064	-0.052	-0.028	-0.070	0.111	0.016	0.142	**0.603**	**0.907**	-0.056	-0.079	0.088	-0.036	0.018
CM	**0.817**	**0.834**	0.114	0.029	0.160*	0.179*	0.179*	0.130*	0.176*	0.139	0.185*	**0.691**	0.163*	0.186*	0.182*	0.174*	0.183*
CA	**0.790**	**0.810**	0.208*	0.107	**0.287**	0.268*	**0.275**	**0.210**	**0.270**	0.175	0.133	**0.601**	**0.259**	**0.289**	**0.259**	**0.256**	**0.265**
CA_c3	0.944*	0.967*	0.832*	0.710*	0.880*	0.883*	0.882*	0.820*	0.874*	0.896*	0.605*	0.172	0.849*	0.882*	0.879*	0.873*	0.877*
NH	**0.898**	**0.911**	0.068	0.017	0.039	0.129*	0.068	0.125*	0.160*	0.200*	**0.473**	**0.838**	0.088	0.039	0.155*	0.065	0.167*
HH	**0.885**	**0.898**	0.149*	-0.016	0.038	0.148*	0.049	**0.209**	0.197*	0.198*	0.403*	**0.821**	0.103*	0.093	**0.205**	0.029	0.193*
Cres	0.290*	**0.324**	**0.231**	0.154*	**0.226**	**0.227**	**0.225**	**0.216**	**0.228**	0.089	0.076	0.201*	**0.221**	0.226*	**0.228**	**0.223**	**0.228**
WBH	0.928*	0.947*	**0.767**	0.588*	**0.830**	**0.830**	**0.829**	**0.756**	**0.822**	0.832*	0.324*	0.924*	**0.795**	**0.830**	**0.825**	**0.817**	**0.824**
EBH	0.829*	0.855*	0.372*	0.115	0.446*	0.461*	0.447*	0.375*	**0.464**	0.380*	0.062	0.726*	0.424*	0.458*	0.456*	0.422*	**0.465**
MR	0.948	0.972	-0.107	-0.578	-0.527	-0.004	-0.305	0.026	0.067	0.300	0.242	0.988	-0.242	-0.260	0.038	-0.364	0.088
SG_c1	**0.895**	**0.913**	**0.914**	**0.871**	**0.930**	**0.931**	**0.930**	**0.909**	**0.928**	**0.932**	**0.858**	**0.942**	**0.918**	**0.930**	**0.930**	**0.927**	**0.929**
SP	**0.732**	**0.756**	**0.453**	0.272*	**0.504**	**0.501**	**0.496**	**0.444**	**0.502**	0.409*	0.061	0.576*	**0.486**	**0.506**	**0.498**	**0.481**	**0.500**
SP_c1	**0.898**	**0.915**	**0.912**	**0.869**	**0.928**	**0.929**	**0.928**	**0.907**	**0.926**	**0.931**	**0.860**	0.942*	**0.916**	**0.928**	**0.928**	**0.925**	**0.927**
CH_c1	**0.875**	**0.890**	**0.909**	**0.873**	**0.922**	**0.923**	**0.923**	**0.905**	**0.921**	**0.921**	**0.861**	**0.921**	**0.913**	**0.923**	**0.922**	**0.920**	**0.922**
SD	**0.919**	**0.924**	0.077*	0.087*	0.113*	**0.137**	0.106*	0.133*	**0.182**	0.163*	**0.605**	**0.861**	0.135*	0.130*	**0.124**	0.062*	**0.179**
MB	**0.911**	**0.917**	0.079*	0.049	0.080*	0.107*	0.061	0.143*	**0.160**	0.123	**0.551**	**0.849**	0.103*	0.121*	0.117*	0.014	0.150*

(B) Table 7B continued from pervious page.

	CB	AB	NS	CM	CA	CA_c3	NH	HH	Cres	WBH	EBH	MR	SG_c1	SP	SP_c1	CH_c1	SD	MB
CB	0.000																	
AB	-0.008	0.000																
NS	-0.063	0.041	0.000															
CM	0.176*	0.184*	0.178*	0.000														
CA	**0.273**	**0.256**	**0.288**	0.043	0.000													
CA_c3	0.889*	0.880*	0.883*	0.602*	0.478*	0.000												
NH	0.059	0.172*	0.047	0.080	0.198	0.790*	0.000											
HH	0.086	0.156	0.070	0.115*	**0.198**	0.768*	0.023	0.000										
Cres	**0.226**	**0.227**	**0.225**	0.166*	0.154*	-0.027	**0.188**	0.184*	0.000									
WBH	**0.837**	**0.826**	**0.831**	**0.465**	0.446*	0.872	**0.716**	0.676*	0.070	0.000								
EBH	0.453*	0.450*	0.458*	0.055	0.092	0.585	0.276*	0.229*	0.035	0.197	0.000							
MR	-0.190	0.064	-0.250	-0.373	-0.266	0.970	-0.484	-0.748	-0.769	0.816	-0.194	0.000						
SG_c1	**0.933**	**0.930**	**0.931**	**0.830**	**0.806**	0.921*	**0.900**	**0.888**	**0.385**	0.916*	**0.843**	0.928	0.000					
SP	**0.502**	**0.489**	**0.505**	0.228*	0.205*	0.454*	**0.401**	0.358*	0.158*	0.012	-0.055	0.098	**0.736**	0.000				
SP_c1	**0.931**	**0.928**	**0.929**	**0.829**	**0.806**	0.923*	**0.898**	**0.887**	**0.379**	**0.919**	0.846*	0.932	-0.030	**0.737**	0.000			
CH_c1	**0.925**	**0.922**	**0.923**	**0.839**	**0.818**	0.901*	**0.897**	**0.888**	**0.455**	**0.902**	**0.849**	0.912	-0.067	**0.754**	0.056	0.000		
SD	0.087*	0.128*	0.127*	**0.248**	**0.311**	0.846*	0.098*	0.129*	**0.429**	**0.785**	**0.442**	-0.144	**0.917**	**0.551**	**0.914**	**0.914**	0.000	
MB	0.063	0.085*	0.105*	**0.202**	**0.257**	0.828*	0.105*	0.059	**0.370**	**0.762**	0.397*	-0.289	**0.909**	**0.500**	**0.906**	**0.906**	0.017	0.000

The resulting maximum likelihood phylogenetic tree created from all *N. californiensis*, *N. gigas*, *N. affinis*, and *U, pugettensis* samples indicated the presence of three cryptic species (Fig. 2). The shrimp samples fell clearly into three distinct and well-supported cryptic species clades, a single *N. gigas* clade from YB, a single *N. affinis* clade, and two *U. pugettensis* clades, one from WGB, GH, and PRE, and one from CH (Fig. 2). While high bootstrap values were generally obtained for established and putative cryptic species clades, internal nodes and those corresponding to relationships within the larger *N. californiensis* population as a whole were poorly resolved.

When samples were partitioned into three putative cryptic species groups and a fourth *N. californiensis* group and then again according to location, the AMOVA indicated that most genetic variation was among these groups (79.61%) but also within populations (15.84%) and among populations within groups (4.55%); each variance component was highly significant ($p < 0.001$; Table 8).

Table 8. AMOVA output for three putative cryptic species groups and one *N. californiensis* group.

Source of variation	*df*	Sum of squares	Variance component	Percentage of variation	p-value ± SE
Among groups	3	2123.787	24.571	79.61	0.000 ± 0.000
Among populations within groups	31	589.466	1.406	4.55	0.000 ± 0.000
Within populations	310	1515.524	4.889	15.84	0.000 ± 0.000

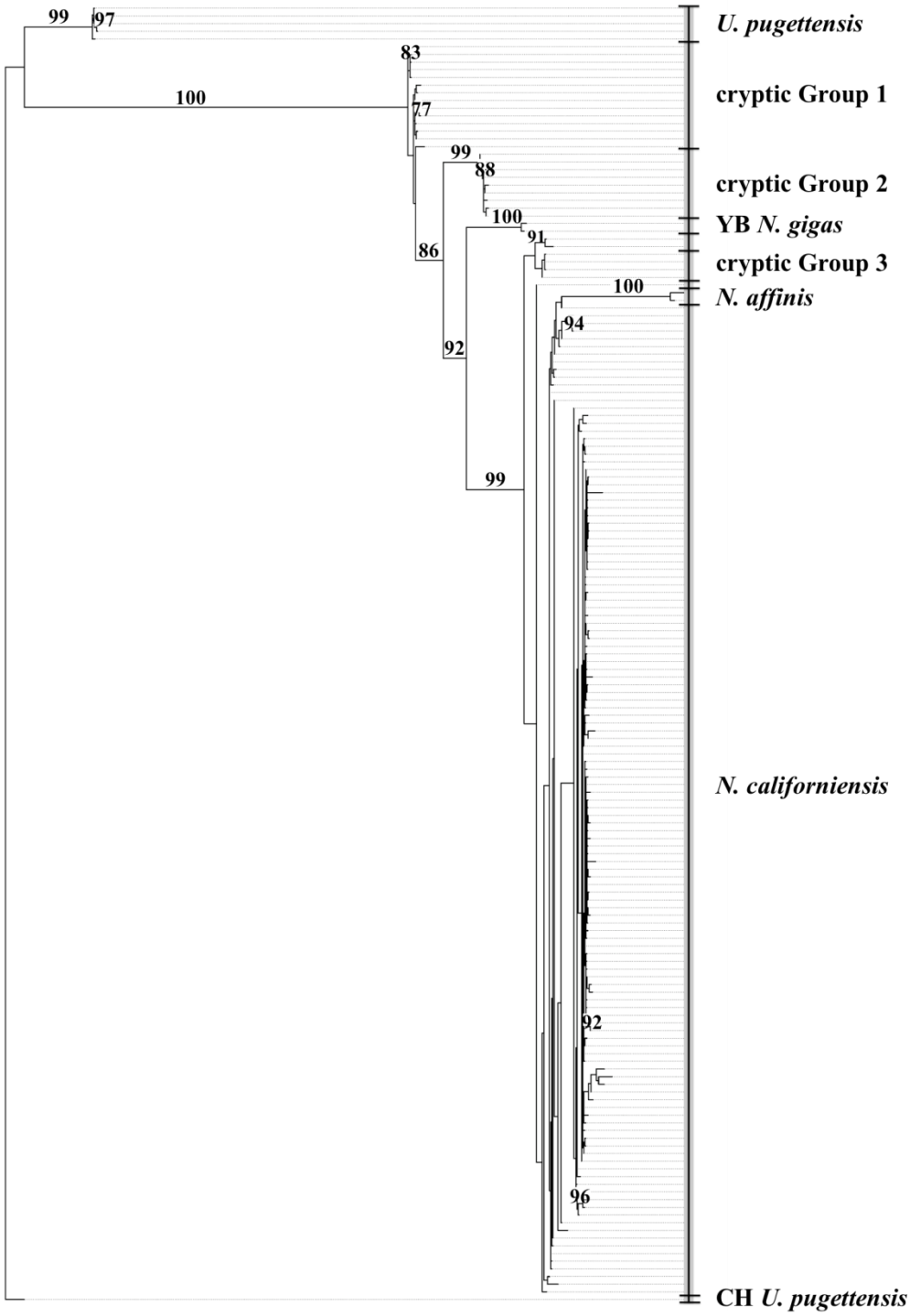
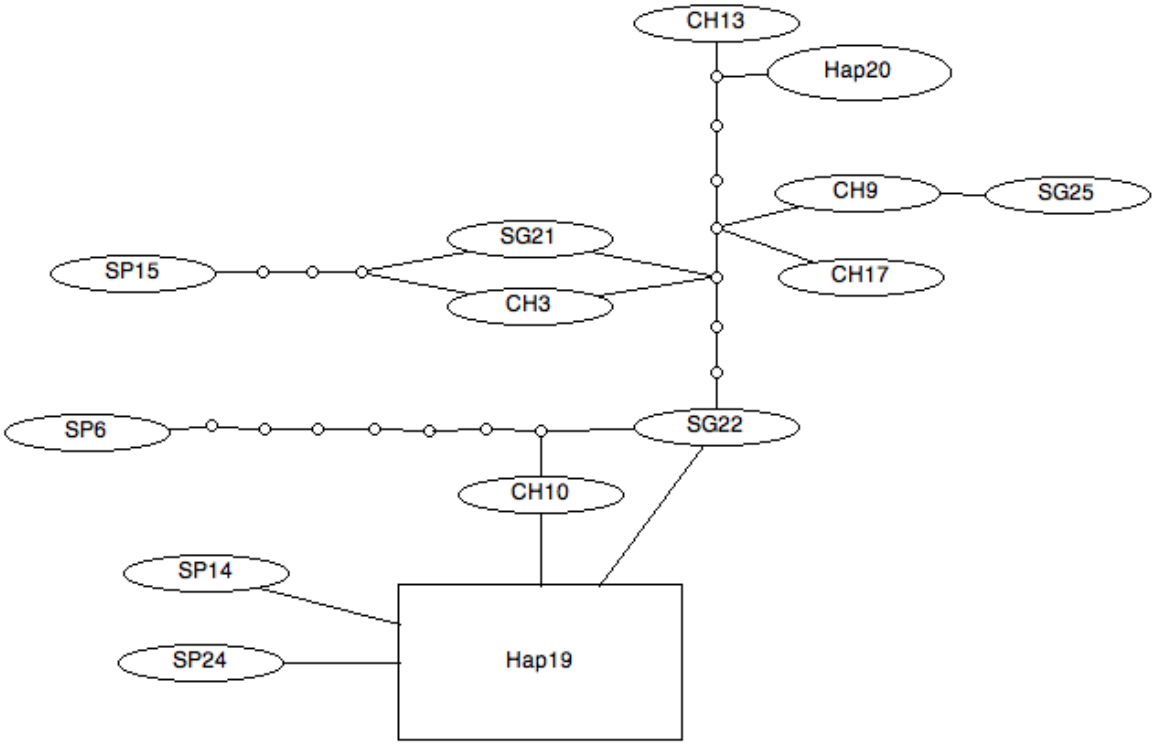


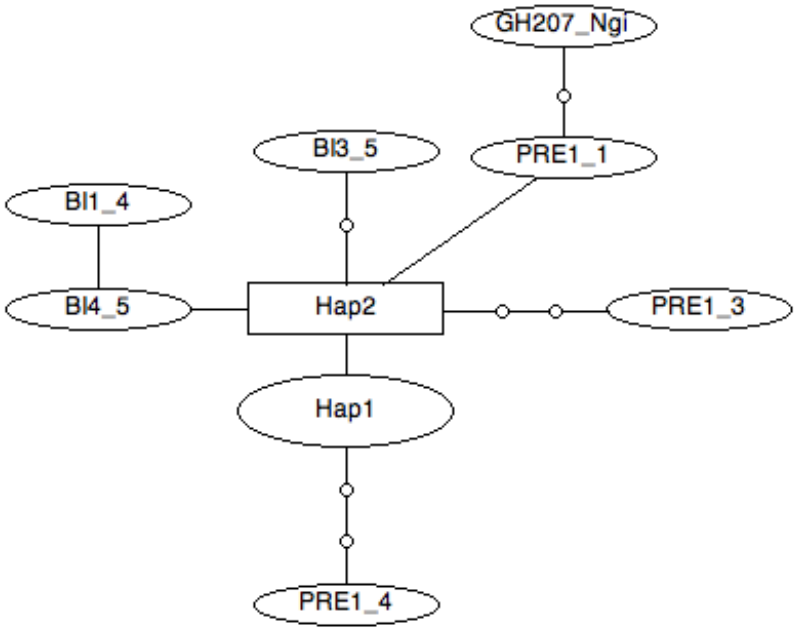
Fig. 2. Maximum likelihood-based phylogenetic tree of *N. californiensis*, *N. gigas*, *N. affinis*, and *U. pugettensis* COI sequences with corresponding bootstrap values; nodes with bootstrap values greater than 75% are shown. Unlabeled sections correspond to *N. californiensis*.

TCS statistical parsimony analysis of all 169 *N. californiensis*, *N. gigas*, *N. affinis*, and *U. pugettensis* haplotypes produced eight unconnected haplotype networks (Fig. 3); four haplotypes could not be connected to any network (CR4065, CM1239, SBB901_Naf, and SBB902_Naf). Furthermore, ten steps, or nucleotide substitutions, were required for parsimonious connections among haplotypes with a confidence limit of 95%. In other words, a haplotype with more than ten nucleotide differences from any other haplotype could not be connected to a network. The main haplotype network that represents *N. californiensis* (data not shown) included two samples recorded as *N. gigas* (GH206_Ngi in Hap2 and ES1_Ngi in Hap4) but did not connect to any *N. affinis* or *U. pugettensis* haplotypes. The haplotype network corresponding to the first putative cryptic species Group 1 (detailed above in Table 3) incorporated all samples from SG and CH and half of the samples from SP (Fig. 3A). The second putative cryptic species Group 2 haplotype network joined all samples from PRE and BI (detailed above in Table 3) as well as two *N. gigas* samples from GH (Fig. 3B). The third putative cryptic species Group 3 haplotype network connected the haplotypes of several CR and CA samples (detailed above in Table 3; Fig. 3C). One unconnected haplotype network linked together all *N. gigas* samples from YB (Fig. 3D) while two others joined all *U. pugettensis* haplotypes from PRE, GH, and WBG (Fig. 3E) and CH (Fig. 3F), respectively. One last unconnected haplotype network included two samples from CA (Fig. 3G).

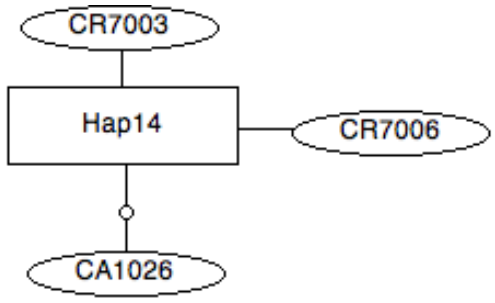
(A)


CH13
Hap20
CH9
SG25
SG21
SP15
CH17
CH3
SP6
SG22
CH10
SP14
SP24
Hap19

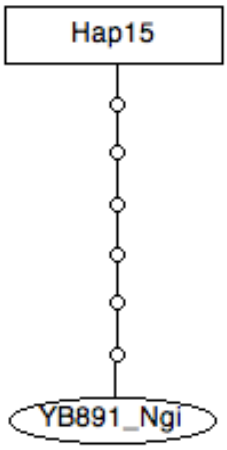
(B)


GH207_Ngi
BI3_5
PRE1_1
BI1_4
BI4_5
Hap2
PRE1_3
Hap1
PRE1_4

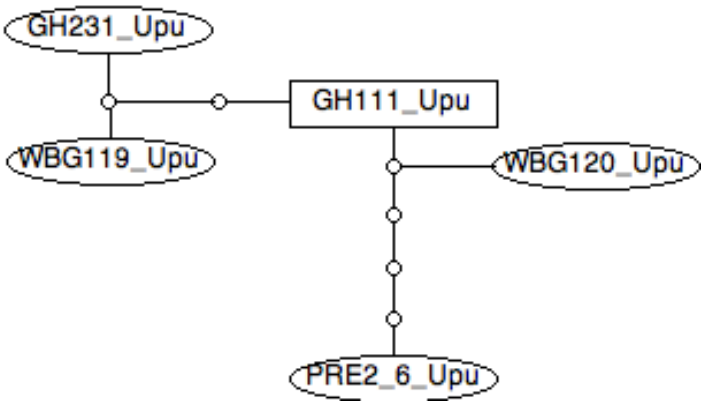
(C)


CR7003
Hap14
CR7006
CA1026

(D)


Hap15
YB891_Ngi

(E)


GH231_Upu
GH111_Upu
WBG119_Upu
WBG120_Upu
PRE2_6_Upu

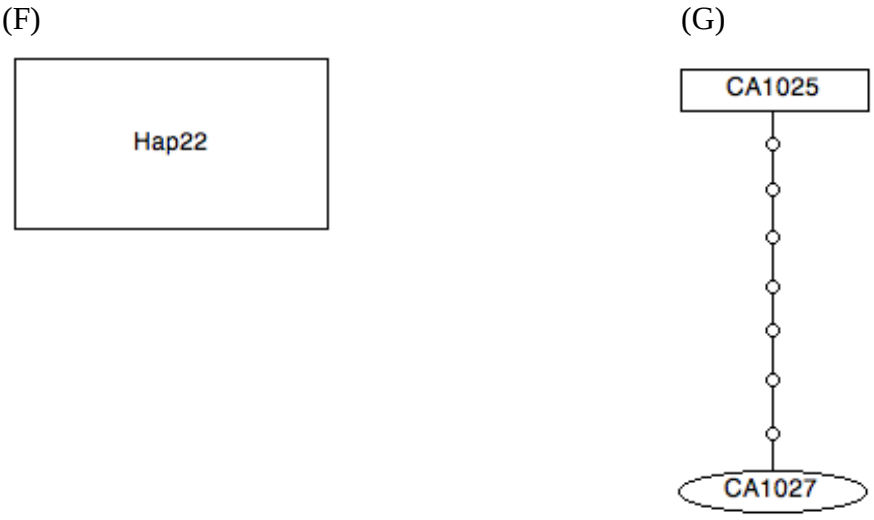


Fig. 3. Putative cryptic species groups, *N. gigas*, and *U. pugettensis* haplotype networks; (A) putative cryptic species Group 1 including samples from SG, CH, and SP; (B) putative cryptic species Group 2 including samples from PRE and BI; (C) putative cryptic species Group 3 including samples from CR and CA; (D) *N. gigas* with samples from YB; (E) *U. pugettensis* with samples from PRE, GH, and WBH; (F) *U. pugettensis* with samples from CH, all with the same haplotype; (G) two connected samples from CH.

Qualitative assessment of eyestalk length revealed that 16 of the 19 specimens (SG28 was excluded because of eyestalks damage) examined from putative cryptic species group 1 had eyestalks that extended to or beyond 1/4 of the length of the second article of the first antennae (Fig. 4A). The eyestalks of the remaining three samples from putative cryptic species group 1 as well as the six samples from the same location (SP) that belonged to the main *N. californiensis* group only protruded as far as the base or less than 1/4 of the length of the second segment of the antenna (Fig. 4B). Concave eyestalk shape (Fig. 4A) was observed in 13 of the 19 putative cryptic species Group 1 specimens and correlated with longer eyestalk length except in the case of three specimens; convex eyestalks (Fig. 4B) were noted in all other Group 1 cryptic species specimens and *N.*

californiensis samples from the same location (SP). The two individuals examined from cryptic species group 2 displayed convex eyestalks that extended to the base of the second article of the first antenna. Measurements of the height and length of the carpus of the major claw were recorded for 14 males in cryptic species group 1 as well as the six *N. californiensis* males from the same location. No quantitative measurements of the male claw were collected for the two available samples from cryptic species group 2. No distinct clusters corresponding to either of the two groups tested were observed when the ratio of the height and length of the carpus of the male major claw was plotted against carapace length (Fig. 5).

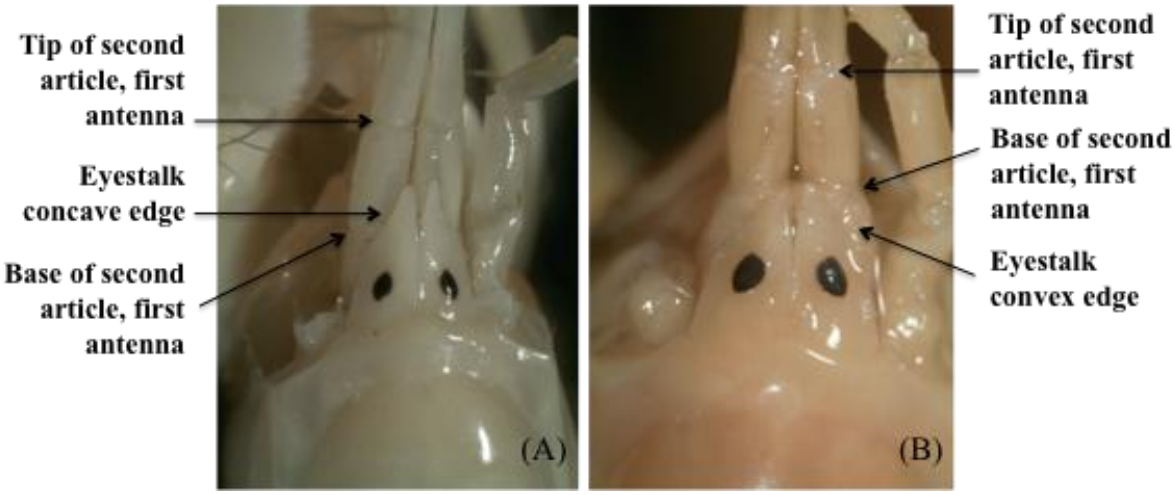


Fig. 4. Dorsal view of shrimp eyestalks and antennae. (A) Representative specimen (SP6) from cryptic species group 1 with concave eyestalks that extend approximately ½ the length of the second article of the first antenna; (B) Representative specimen (SP8) from the large *N. californiensis* group with convex eyestalks that extend to the base of the second article of the first antenna.

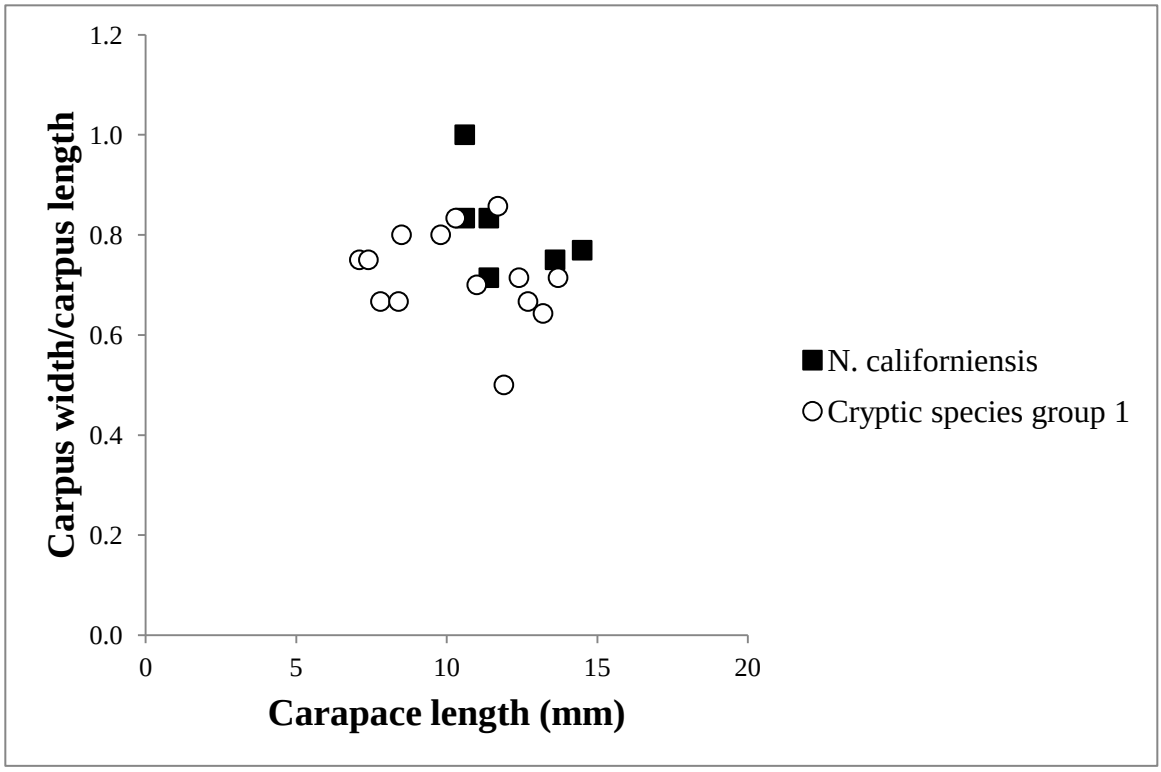


Fig. 5. Patterns of morphological variation in the major claw as compared to the carapace in *N. californiensis* and cryptic species Group 1 males.

DISCUSSION

Larval Dispersal and Population Structure. Much debate presently surrounds the dispersal potential of *N. californiensis* larvae. Analyses of *N. californiensis* larvae and adults in Washington, Oregon, and California have shown little genetic differentiation among localities (Kozuka 2008; Pernet et al. 2008; Buncic 2010), a result that points to extensive gene flow associated with long distance larval transport in the open ocean. Conversely, examinations of *N. californiensis* offshore distribution suggest that larvae moderate swimming behavior to remain in shallow coastal waters (Morgan et al. 2009). The study presented here provided further genetic evidence that indicates that

N. californiensis larvae found along the west coast of the United States generally travel great distances during the pelagic dispersal phase; however, select populations may be locally retained.

Observations of mtDNA sequence differences revealed high haplotype diversity (h = 0.877) as well as a large number of unique haplotypes; approximately 78% of the total haplotypes occurred only once. However, overall nucleotide diversity was low (π = 0.017) indicating that nucleotide sequences were generally similar among haplotypes. Despite the large number of singletons, the two most common haplotypes (Hap4 and Hap5) were found in Washington, Oregon, and California, suggesting that historic or contemporary gene flow, and thus larval dispersal, is likely to occur extensively along the U.S. Pacific coast. The same pattern of high haplotype diversity coupled with low nucleotide diversity has been detected in several other marine invertebrate species with long-distance planktonic larval dispersal including the mud crab (*Scylla serrata*; Fratini and Vannini 2002), sea cucumber (*Holothuria nobilis*; Uthicke and Benzie 2003), the lined shore crab (*Pachygrapsus crassipes*; Cassone and Boulding 2006), and two gastropods (*Litorinna keenae* and *Concholepas concholepas*; Lee and Boulding 2007; Cardenas et al. 2009). The hypothesis of long-distance *N. californiensis* larval dispersal is further supported by the results of a Mantel test that revealed no correlation between genetic and geographic distance and AMOVA that showed that only 16% of genetic variation was explained by locality (once putative cryptic specimens had been separately analyzed); the presence of isolation by distance would have been evidence for constrained dispersal.

The high number of mtDNA haplotypes (disproportionate to the generally low divergence seen between haplotypes) is mirrored in large, negative, and statistically significant estimates of Fu's F_S. The values of Fu's F_S are also consistent with an excess of low-frequency variants, recent population expansion, and implies that selection is present at this locus (Ray et al. 2003). The Tajima's *D* statistic is much less powerful and sensitive to demographic expansion than Fu's F_S (Fu 1997; Ramos-Onsins & Rozas 2002; Ray et al. 2003; Wares 2011) so it was not surprising that values of Tajima's *D* were mixed positive and negative and non-significant. Fu's F_S is also the most appropriate test of neutrality for the detection of population growth for large sample sizes (Ramos-Onsins & Rozas 2002).

It is generally held that marine species that experience a long larval life stage will disperse widely and show weak genetic population structure (Burton 1983; Bohonak 1999; Kelly & Palumbi 2010). Interestingly, high levels of F_{ST} were found among adult *N. californiensis* populations indicating extreme genetic subdivision among sampling locations. Several potential scenarios explain why estimates of F_{ST} seem to contradict expectation as well as previously presented haplotype diversity, Mantel test, and AMOVA data. An immediately obvious possibility is that the observed genetic differentiation among populations is merely an artifact of performing F_{ST} analyses on a group of samples with such a high number of unique haplotypes (Frantini & Vannini 2002). The non-random distribution of haplotypes could also be caused by the presence of an *N. californiensis* metapopulation. A metapopulation that spans the entire west coast of the United States would display chaotic local variation that would not be correlated

with geographic region, much like that which was detected here. Sea urchin species (*Tripneustes* spp.) found worldwide that also utilize pelagic larval dispersal are often structured as large metapopulations (Lessios et al. 2003). Finally, the potential for the seemingly contrasting outcomes of shared haplotypes among states and strong population structure to have been caused by the transport of *N. californiensis* across regional boundaries as live bait cannot be ruled out (Pernet et al. 2008). The observed patterns of genetic differentiation could also be achieved if *N. californiensis* larvae are naturally locally retained but regularly imported across great distances and into foreign estuaries by anglers (discussed in Pernet et al. 2008).

Perhaps the most plausible reason for these seemingly inconsistent results is that historic gene flow and high dispersal did occur in these regions (and may still occasionally), thus accounting for shared haplotypes among Washington, Oregon, and California. However, changing ocean conditions and increased habitat heterogeneity now inhibit successful recruitment of larvae with long-distance dispersal potential, which has resulted in extreme adult population structure (Burton & Feldman 1981). Thermal stress has been shown to induce a genetic signature in acorn barnacle populations (*Semibalanus balanoides*; Bertness & Gaines 1993) while wave condition and hydrodynamic stress is known to cause mortality and affect the subsequent degree of larval dispersal and genetic population structure in mussels (*Perna perna*; Nicastro et al. 2008). The many ways for selection to trigger local adaptation and genetic differentiation among local populations of marine species with widely dispersing larvae have been documented but are variable and not well understood (Palumbi & Kelly 2010).

It is also important to note that while the results of this investigation also support a hypothesis of wide-ranging *N. californiensis* larval dispersal, details are not consistent with earlier studies. Kozuka (2008) and Buncic (2010) examined genetic diversity at the same COI locus in *N. californiensis* larvae in Oregon and Washington and found high haplotype and low nucleotide diversity (h = 0.958, π = 0.045, Kozuka 2008; h = 0.952, π = 0.040 in 2005 and h = 0.984, π = 0.037 in 2006, Buncic 2010) much like the adult populations examined here; results of neutrality tests and large negative estimates of Fu's F_S were likewise similar. However, F_{ST} values presented by both Kozuka (2008; ranging from 0.051 to 0.483) and Buncic (2010; ranging from 0.057 to 0.620) generally show weak structure among populations that pale in comparison to F_{ST} levels witnessed in adults (Table 7). Disparate accounts of population structure between larvae and adult *N. californiensis* populations could conceivably be explained by the stochastic nature of larval recruitment and establishment into estuaries. In other words, just because larvae are sampled offshore at a particular location does not mean that they will settle nearby or survive past the larval stage. While individual larvae may be transported far from the natal estuary and mix extensively with other larvae of the same year class, it is likely those that arrive at and are best suited to a particular local micro-environment will thrive and remain there, resulting in significant genetic divergence among localities (Grosberg & Cunningham 2001).

Results of genetic population structure analyses by AMOVA of adult *N. californiensis* performed here are comparable to those obtained by Pernet et al. (2008). Pernet et al. (2008) collected adult *N. californiensis* specimens from throughout

California and from one site in Oregon and one in Washington and found that 29% of molecular variation was associated with geographic positioning either north or south of Point Conception, a known marine biogeographic boundary in California (Dawson et al. 2011). After removing individual specimens thought to belong to a putative cryptic clade ("clade A") found in southern California baitshops and estuaries (Carpinteria Marsh, Marina Del Rey, San Pedro, Alamitos Bay, Anaheim Bay, Agua Hedionda Lagoon, and San Diego), Pernet et al. (2008) repeated the AMOVA and determined that 99% of molecular variation was explained by local population affiliation, suggesting that significant population structure existed among sampling sites that was not caused by the Pt. Conception boundary. Because Pt. Conception was not initially found to be a major cause of population structure in either report, a similar AMOVA was not run here. Instead, specimens were divided by cryptic species grouping, a categorization strategy that explained the majority (79.61%) of genetic variation. Such a great percentage of explained variation suggests that in both Pernet et al. (2008) and this study, cryptic speciation is likely to be the cause of extreme genetic structure. Furthermore, putative species Group 1 is likely the same as "clade A" from Pernet et al. (2008). Both groups are found in southern California estuaries and appear sympatric with the larger *N. californiensis* population.

One obvious remaining quandary is the cause of incongruence between distribution surveys that show local larval retention and genetic studies that point to long-distance larval dispersal potential. It is conceivable that these experimental results conflict for many of the same reasons that haplotype diversity and pairwise F_{ST} values

seemingly do not agree. As previously described, the transport of live marine species bait across regional boundaries for recreational fishing has been well documented (Ludwig & Leitch 1996; Cohen et al. 2001; Weigle et al. 2005) and would allow both scenarios to occur simultaneously. It is widely held that as little as one migrant or translocated individual per generation is sufficient to cause a smoothening of genetic differentiation and prevent the complete fixation of alleles among sampling sites (Wright 1931; Kimura and Ohta 1971; Lewontin 1974; Spieth 1974) which could also explain why the same mitochondrial haplotypes are shared among shrimp in distant estuaries. *N. californiensis* larvae could also alter their behavior to remain within or nearby natal estuaries and avoid offshore advection (Kingsford et al. 2002) while anglers could transfer adults into foreign estuaries via bait buckets thus facilitating dispersal along the west coast of the U.S. It is again possible that historic gene flow may have occurred across a great distance but habitat heterogeneity in the form of changing currents, salinity, or temperature over time has encouraged contemporary local adaptation to new environmental conditions and thus local retention of *N. californiensis* larvae (Sanford & Kelly 2011). Swimming behaviors may also be inconsistent among discrete shrimp populations and could change according to stochastic events or annual variation causing sporadic local retention seen in distribution studies. Finally, it is important to consider that larvae found nearshore in distribution studies may not ever reach adulthood or recruit into the population. An examination of larval distribution and genetics in parallel with adult genetics among estuaries across multiple years would paint a more complete picture of these processes.

Cryptic species analysis. The second facet of this study involved the identification and evaluation of cryptic species found in estuaries along the western coast of the U.S. Multiple lines of evidence including a phylogenetic tree constructed using maximum likelihood, the partitioning of variance by AMOVA, and haplotype networks clearly support the existence of three putative cryptic species groups among sampled individuals. The largest cryptic species Group 1 was found in SG, SP, and CH and is likely to be the same as that identified by Pernet et al. (2008), known as "clade A"; cryptic species Groups 2 (found in PRE, BI, and GH in Washington) and 3 (found in CR and CA in Washington and Oregon respectively) are unique to this study. However, no genetic test or morphological measurement performed here was able to definitively diagnose members of putative cryptic groups as *N. gigas* or a presently unknown species.

Presently, four species of burrowing shrimp (two of which are ghost shrimp) are known to occur in North American west coast estuaries: *N. californiensis* and its sympatric, larger, but often indistinguishable closest relative *N. gigas* (both ghost shrimp; Stevens 1928; Pernet et al. 2010), *N. affinis*, and *Upogebia pugettensis* (A. F. D'Andrea, pers. comm.). If cryptic species Groups 1, 2, and 3 were simply populations of *N. gigas*, it is expected that all specimens would display the morphological features characteristic of the species such as elongated and concave eyestalks and larger length to width ratio of the carpus of the male claw. However in cryptic species Group 1 (the only group with enough remaining in tact specimens to perform and draw inferences from morphological analyses) only 84% of individuals possessed eyestalks whose length was indicative of *N. gigas* and only 68% of individuals had the typical concave *N. gigas* eyestalk shape; major

claw measurements were uninformative. It is possible that the cryptic species groups identified in this study are indeed members of an *N. gigas* population that has hybridized with *N. californiensis* at some point in the past. Such an event could results in identical or very similar maternally inherited mtDNA haplotypes but mixed autosomal morphological traits. Bernatchez et al. (1995) observed a similar pattern of introgression in fish is Lake Alain, Québec, Canada, that possessed an mtDNA haplotype that matched that of the Québec Arctic char (*S. alpinus*) but were morphologically indistinguishable from brook trout (*S. fontinalis*).

Haplotype and phylogenetic evidence using this segment of COI was similarly unable to conclusively resolve whether the three putative cryptic species were undiagnosed *N. gigas* clades. Cryptic species Groups 1 and 3 contained no specimens identified as *N. gigas* at the time of collection and were present as distinct and unique clades, suggesting that these clustering of individuals may indeed represent two unrecognized species. Cryptic species group 2 included two individuals from GH categorized as *N. gigas* at the time of collection: one with a private haplotype and one that displayed Hap2 that was shared by other putative *N. californiensis* shrimp in the same group. Several scenarios potentially explain these findings. Given that it is often difficult to distinguish *N. californiensis* from *N. gigas*, it is possible that the specimens collected from GH were incorrectly identified as *N. gigas* during the sampling process or that the mixed morphological traits described above are characteristic of this cryptic species group or represent a past hybridization event. Interestingly, the single *N. gigas* specimen from ES exhibited Hap4 that was shared by several other *N. californiensis*

individuals that grouped with large *N. californiensis* clade. This single anomaly may again point to the fallibility of morphological measurements when taken alone for species identification at the time of collection. While it is difficult to make inferences about species identification based on available morphological, haplotype, and phylogenetic data, none of these clades grouped with *N. gigas* individuals from YB which lends support to the idea that they may indeed be unknown and distinct species.

The interaction between *N. californiensis*, *N. gigas*, and putative unknown species groups in eastern Pacific estuaries is likely to be complex given the inconsistent nature of morphological and genetic data. It could be that, unlike *N. californiensis*, *N. gigas* dispersal is highly constrained, operating across very short distances only. Such a mode of dispersal taking place over a long period of time would result in genetically divergent *N. gigas* populations, with potentially unique local adaptations that might appear to be unique species. It is also possible that reproductively isolated populations of *N. californiensis* or *N. gigas* that cannot reach the open ocean for pelagic dispersal were introduced to and settled in foreign estuaries, also manifesting as putative cryptic species. The potential biological explanations for the results surrounding cryptic speciation obtained in this study are numerous and varied. Future studies of genetic structure and cryptic speciation should focus on the analysis of multiple genetic markers including additional mtDNA or nuclear gene regions or microsatellites to improve resolution and detection of relationships among populations of ghost shrimp.

LITERATURE CITED

Alexander SE, Roughgarden J (1996) Larval transport and population dynamics of intertidal barnacles: a coupled benthic model. Ecol Monogr 66:259-275

Baker P (1995) Review of ecology and fishery of the Olympia oyster, Ostrea lurida with annotated bibliography. J Shellfish Res 14:501-518

Bernatchez L, Glemet H, Wilson CC, Danzmann RG (1995) Introgression and fixation of Arctic char (*Salvelinus alpinus*) mitochondrial genome in allopatric population of brook trout (*Salvelinus fontinalis*). Can J Fish Aquat Sci 52:179-185

Bertness MD, Gaines SD (1993) Larval dispersal and local adaptation in acorn barnacles. Evolution 47:316-320

Bird EM (1982) Population dynamics of thalassinidean shrimps and community effects through sediment modification. PhD dissertation, University of Maryland, College Park, Maryland

Bohonak AJ (1999) Dispersal, gene flow, and population structure. Q Rev Biol 74:21-45

Buncic M (2010) Interannual differences in the estuarine ghost shrimp, *Neotrypaea californiensis*. MS thesis, San Jose State University, San Jose, California

Burton RS (1983) Protein polymorphisms and genetic differentiation of marine invertebrate populations. Mar Biol Lett 4:193-206

Burton RS, Feldman MW (1981) Population genetics of *Tigriopus californicus* II. Differentiation among neighboring populations. Evolution 35:119-1205

Cardenas L, Castilla JC, Viard F (2009) A phylogeographical analysis across three biogeographical provinces of the south-eastern Pacific: the case of the marine gastropod *Concholepas concholepas*. J Biogeogr 36:969-981

Cassone BJ, Boulding EG (2006) Genetic structure and phylogeography of the lined shore crab, *Pachygrapsus crassipes*, along the northeastern and western Pacific Coasts. Mar Biol 149:213-226

Chew KK (1984) Recent advances in the cultivation of molluscs in the Pacific United States and Canada. Aquaculture 39:69-81

Clement M, Posada D, Crandall KA (2000) TCS: a program to estimate gene genealogies. Mol Ecol 9:1657-1660

Cohen AN, Weinstein A, Emmett MA, Lau W, Carlton JT (2001) Investigations into the introduction of non-indigenous marine organisms via the cross-continental trade in marine baitworms. Report for U.S. Fish and Wildlife Service, San Francisco Bay Program

Cowen RK (2002) Larval dispersal and retention and consequences for population connectivity. In: Sale PF (ed) Coral reef fishes, dynamics and diversity in a complex ecosystem. Academic, San Diego, California

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Annu Rev Mar Sci 1:443-466

Dawson MN (2001) Phylogeography in coastal marine animals: a solution from California? J Biogeogr 28:723-736

Dorman CE, Dever EP, Largier JL (2005) Buoy measured wind stress and curl of the wind stress over the shelf off Bodega Bay, California. Deep-Sea Res Part II 53:2850-2864

Dumbauld BR, Armstrong DA, Feldman KL (1996) Life-history characteristics of two sympatric Thalassinidean shrimps, *Neotrypaea californiensis* and *Upogebia pugettensis*, with implications for oyster culture. J Crustac Biol 16:689-708

Dumbauld BR, Booth S, Cheney D, Suhrbier A, Beltran H (2006) An integrated pest management program for burrowing shrimp control in oyster aquaculture. Aquaculture 261:976-992

Ersts PJ (2012) Geographic Distance Matrix Generator (version 1.2.3), American Museum of Natural History, Center for Biodiversity and Conservation. Accessed 17 June. http://biodiversityinformatics.amnh.org/open_source/gdmg

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47-50

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491

Farrell TMD, Bracher D, Roughgarden J (1991) Cross-shelf transport causes recruitment to intertidal populations in central California. Limnol Oceanogr 36:279-288

Feldman KL, Armstrong DA, Dumbauld BR, DeWitt TH, Doty DC (2000) Oysters, crabs, and burrowing shrimp: review of an environmental conflict over aquatic resources and pesticide use in Washington state's (USA) coastal estuaries. Estuaries 23:141-176

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368-376

Ferretti L, Perez-Enciso M, Ramos-Onsins S (2010) Optimal neutrality tests based on the frequency spectrum. Genetics 186:353-365

Ford MJ (2002) Applications of selective neutrality tests to molecular ecology. Mol Ecol 11:1245-1262

Fratini S, Vannini M (2002) Genetic differentiation in the mud crab *Scylla serrata* (Decapoda: Portunidae) within the Indian Ocean. J Exp Mar Biol Ecol 272:103-116.

Fu Y (1997) Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925

Glez-Peña D, Gomez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D (2010) ALTER: program-oriented conversion of DNA and protein alignments. Nucleic Acids Res 38:W14-W18

Grosberg RK, Cunningham CW (2001) Genetics structure in the sea: from populations to communities. In: Bertness MD, Hay ME, Gaines SD (eds) Marine community ecology. Sinauer, Sunderland, Massachusetts

Hart JFL (1982) Crabs and their relatives of British Columbia. Br Columbia Prov Mus Handb 40, Victoria, British Columbia

Hudson RR (1990) Gene genealogies and the coalescent process. In: Futuyama D, Antonovics J (eds) Oxford surveys in evolutionary biology, Vol. 7. Oxford University Press, New York, New York

Huelsenbeck JP (1995) Performance of phylogenetic methods in simulation. Syst Biol 44:17-48

Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276:227-232

Johnson GE, Gonor JJ (1982) The tidal exchange of *Callianassa californiensis* (Crustacea, Decapoda) larvae between the ocean and the Salmon River estuary, Oregon. Estuar Coast Shelf Sci 14:501-516

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the Northeastern Pacific rocky intertidal community. PLoS One 5:e8594

Kimura M, Ohta T (1971) Theoretical aspects of population genetics. Princeton University Press, Princeton, New Jersey

Kingsford MJ, Leis JM, Shanks A, Lindeman KC, Morgan SG, Pineda J (2002) Sensory environments, larval abilities and local self-recruitment. Bull Mar Sci 70:309-340

Koracin D, Dorman CE, Dever EP (2004) Coastal perturbations of marine-layer winds, wind stress, and wind stress curl along California and Baja California in June 1999. J Phys Oceanogr 34:1152-1173

Kozuka K (2008) Dispersal and population structure of *Neotrypaea californiensis*. MS thesis, San Jose State University, San Jose, California

Kuhner M, Felsenstein J (1994) A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rate. Mol Biol Evol 11:459-468

Largier J (2003) Considerations in estimating larval dispersal distances from oceanographic data. Ecol Appl 13:S71-89

Larkin M (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948

Lee HJ, Boulding EG (2007) Mitochondrial DNA variation in space and time in the northeastern Pacific gastropod, *Littorina keenae*. Mol Ecol 16:3084-3103

Leis JM (2006) Are larvae of demersal fishes plankton or nekton? Adv Mar Biol 51:59-141

Leis JM, McCormick ML (2002) The biology, behavior and ecology of the pelagic, larval stage of coral reef fishes. In: Sale PF (ed) Coral reef fishes, dynamics and diversity in a complex ecosystem. Academic, San Diego, California

Lessios HA, Kane J, Robertson DR (2003) Phylogeography of the pantropical sea urchin *tripneustes*: contrasting patterns of population structure between oceans. Evolution 57:2026-2036

Lewontin RC (1974) The genetic basis of evolutionary change. Columbia University Press, New York, New York

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

Lipphardt B, Small D, Kirwin A, Wiggins S, Ide K, Grosch CE, Paduan JD (2006) Synoptic Lagrangian maps: applications to surface transport in Monterey Bay. J Mar Res 64:221-247

Ludwig HR, Leitch JA (1996) Interbasin transfer of aquatic biota via angler's bait buckets. Fisheries 21:14-18

Mace AJ, Morgan SG (2006) Biological and physical coupling in the lee of a small headland: contrasting transport mechanisms for crab larvae in an upwelling region. Mar Ecol Prog Ser 324:185-196

MacGinitie GE (1934) The natural history of *Callianassa californiensis* Dana. Am Midl Nat 15:166-177

Manning RB, Felder DL (1991) Revision of the American Callianassidae (Crustacea: Decapoda: Thalassinidea). Proc Biol Soc Wash 104(4):764-792

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209-220

McCrow LT (1972) The ghost shrimp, *Callianassa californiensis* Dana, 1854, in Yaquina Bay, Oregon. MS thesis, Oregon State University, Corvallis, Oregon

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Mol Ecol 11:5-18

Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T (2010) The CIPRES Portals. CIPRES. Accessed 17 June. http://www.phylo.org/sub_sections/portal

Morgan SG, Fisher JL (2010) Larval behavior regulates nearshore retention and offshore migration in an upwelling shadow and along the open coast. Mar Ecol Prog Ser 404:109-126

Morgan SG, Fisher JL, Miller SH, McAfee ST, Largier JL (2009) Nearshore larval retention in a region of strong upwelling and recruitment limitation. Ecology 90:3489-3502

Nicastro KR, Zardi GI, McQuaid CD, Teske PR, Barker NP (2008) Coastal topography drives genetic structure in marine mussels. Mar Ecol Prog Ser 368:189-195

Parrish RH, Nelson CS, Bakun A (1981) Transport mechanisms and the reproductive success of fishes in the California Current. Biol Oceanogr 1:175-203

Pernet B, Deconinck A, Llaban A, Archie JW (2008) Evaluating risks associated with transport of the ghost shrimp *Neotrypaea californiensis* as live bait. Mar Biol 153:1127-1140

Pernet B, Deconinck A, Haney L (2010) Molecular and morphological markers for distinguishing the sympatric intertidal ghost shrimp, *Neotrypaea californiensis* and *N. gigas* in the eastern Pacific. J Crustac Biol 30:323-331

Pimentel GE (1983) Recruitment of larvae of the ghost shrimp, *Callianassa californiensis* (Crustacea, Decapoda) in the South Slough estuary, Oregon. MS thesis, Oregon State University, Corvallis, Oregon

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. Trends Ecol Evol 16:37-45

Posey MH (1986a) Changes in a benthic community associated with dense beds of a burrowing deposit feeder, *Callianassa californiensis*. Mar Ecol Prog Ser 31:15-22

Posey MH (1986b) Predation on burrowing shrimp: distribution and community consequences. J Exp Mar Biol Ecol 103:143-161

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Mol Biol Evol 19:2092-2100

Ray N, Currant M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol 20:76-86

Roughgarden J, Gaines S, Possingham HP (1988) Recruitment dynamics in complex life cycles. Science 241:1460-1466

Roughgarden J, Pennington JT, Stoner D, Alexander S, Miller K (1992) Collisions of upwelling fronts with the intertidal zone: the cause of recruitment pulses in barnacle populations of central California. Acta Oecol 12:35-51

Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. Annu Rev Mar Sci 3:509-535

Shanks AL (2009) Pelagic larval duration and dispersal distance. Biol Bull 216:373-385

Shanks AL, Eckert GL (2005) Population persistence of California Current fishes and benthic crustaceans: a marine drift paradox. Ecol Monogr 75:505-524

Shanks AL, Granthan BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. Ecol Appl 13:S159-S169

Shumway SE (2011) Shellfish aquaculture and the environment. Wiley Blackwell Science Publishers, Ames, Iowa

Smith AE, Chapman JW, Dumbauld BR (2008) Population structure and energetics of the bopyrid isopod parasite *Orthione griffensis* in mud shrimp *Upogebia pugettensis*. J Crustac Biol 28:228-233

Spieth PT (1974) Gene flow and genetic differentiation. Genetics 78:961-965

Sponaugle S, Cowen RK, Shanks A, Morgan SG, Leis JM, Pineda J, Boehlert GW, Kingsford MJ, Lindeman KC, Grimes C, Munro JL (2002) Predicting self-recruitment in marine populations biophysical correlates and mechanisms. Bull Mar Sci 70:341-375

Sponaugle S, Lee T, Kourafalu V, Pinkard D (2005) Florida Current frontal eddies and the settlement of coral reef fishes. Limnol Oceanogr 50:1033-1048

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Stevens BA (1928) Callianassidae from the west coast of North America. Publications – Puget Sound Biological Station 6:315-369

Stevens BA (1929) Ecological observations on Callianassidae of Puget Sound. Ecology 10:399-405

Strub PT, Allen JS, Huyer A, Smith RL (1987) Seasonal cycles in currents, temperature, winds, and sea level over the Northeast Pacific continental shelf: 35° N to 48° S. J Geophys Res 92:1507-1526

Strub PT, James C (2000) Altimeter-derived variability of surface velocities in the California Current System: 2. Seasonal circulation and eddy statistics. Deep-Sea Res Part II 47:831-870

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595

Tateno Y, Takezaki N, Nei M (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum parsimony methods when substitution rate varies with site. Mol Biol Evol 11:261-267

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associationswith haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619-633

U.S. Department of Agriculture, National Agriculture Statistics Service (2006) 2002 Census of agriculture: Census of aquaculture (2005). United States Department of Agriculture, Volume 3, Special Studies Part 2, AC-02-SP-2. Beltsville, Maryland

Uthicke S, Benzie JAH (2003) Gene flow and population history in high dispersal marine invertebrates: mitochondrial SNA analysis of *Holothuria nobilis* (Echinodermata: Holothuroidea) populations from the Indo-Pacific. Mol Ecol 12:2635-2648

Wares JP (2011) Insights from population genetics: are all crustaceans created equal? J Crustac Biol 31:339-344

Washington Department of Fish & Wildlife (1970) Ghost shrimp control experiments with Sevin, 1960-1968. Washington Department of Fisheries Technical Report 1, Olympia, Washington

Weigle SM, Smith LD, Carlton JT, Pederson J (2005) Assessing the risk of introducing exotic species via the live marine species trade. Conserv Biol 19:213-223

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97-259

Wright S (1943) Isolation by distance. Genetics 28:114-138

Wright S (1951) The genetical structure of populations. Ann Eugenic 1:323-334

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395-420

Yang Z, Rannala B (2012) Molecular phylogenetics: principles and practice. Nat Rev Genet 13:303-314

Yoshioka PM (1982) Role of planktonic and benthic factors in the population dynamics of the bryzoan, *Membranipora membranacea*. Ecology 63:457-468