Spring 2015

Investigation of Population Structure and Distribution of the Invasive Bryozoan Watersipora Species along the California Coast using Nuclear and Mitochondrial DNA

Darren James Wostenberg
San Jose State University

INVESTIGATION OF POPULATION STRUCTURE AND DISTRIBUTION OF THE INVASIVE BRYOZOAN *WATERSIPORA* SPECIES ALONG THE CALIFORNIA COAST USING NUCLEAR AND MITOCHONDRIAL DNA

A Thesis Presented to

The Department of Biological Sciences

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Darren J. Wostenberg

May 2015

The Designated Thesis Committee Approves the Thesis Titled

INVESTIGATION OF POPULATION STRUCTURE AND DISTRIBUTION OF THE INVASIVE BRYOZOAN *WATERSIPORA* SPECIES ALONG THE CALIFORNIA COAST USING NUCLEAR AND MITOCHONDRIAL DNA

by

Darren J. Wostenberg

APPROVED FOR THE DEPARTMENT OF BIOLOGICAL SCIENCES

SAN JOSÉ STATE UNIVERSITY

May 2015

Dr. Leslee Parr	Department of Biological Sciences, San José State University
Dr. Joshua Mackie	Department of Biological Sciences, San José State University
Dr. Sean Craig	Department of Biological Sciences, Humboldt State University

ABSTRACT

INVESTIGATION OF POPULATION STRUCTURE AND DISTRIBUTION OF THE INVASIVE BRYOZOAN *WATERSIPORA* SPECIES ALONG THE CALIFORNIA COAST USING NUCLEAR AND MITOCHONDRIAL DNA

by Darren J. Wostenberg

This study combined microsatellite nuclear DNA analysis with cytochrome *c* oxidase 1 (COI) mitochondrial DNA analysis to evaluate coastal population structure, environmental factors influencing population distribution, and the potential for hybridization among coexisting *Watersipora* haplogroups along the California coast. Mitochondrial DNA analysis of the COI gene identified three haplogroups: *W. subtorquata* clade A, *W. subtorquata* clade B, and *W.* new species. Analyses resulted in seven haplotypes for haplogroup clade A, and a single haplotype in each haplogroup clade B and new species. Microsatellite data indicated the greatest source of genetic variation in the two species examined (*W. subtorquata* and *W.* new species) was within individuals of the population (53.7% and 69.3%, respectively), compared to among individuals (36.2% and 20.2%, respectively), populations (5.5% and 10.5%, respectively), and regions (4.6% and 0.0%, respectively). Congruence analysis between mitochondrial and nuclear data correctly matched nuclear genotypes with mitochondrial haplogroups. Evidence of hybridization was not detected among the two *Watersipora* species, notwithstanding one highly variable locus. Points of introduction could not be identified; however, locations in regions with high ship traffic displayed a greater number of total alleles.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my wife, Jacquelyn Wostenberg, for her years of patience, encouragement, and support. I could not have completed this journey without her, and her support is greatly appreciated. I would also like to thank my parents, Bill and Melody Wallis, and my brother, Jeremy Wostenberg, for their constant support throughout my career as a graduate student. I would like to extend my deepest thanks and respect to my thesis committee: Dr. Leslee Parr, Dr. Joshua Mackie, and Dr. Sean Craig. The guidance and support I have received from each of you has been beneficial and rewarding beyond words. Thank you, Dr. Parr for accepting me as a graduate student, for guiding me through graduate school, and for your professional support. Thank you, Dr. Mackie for guiding my research, for the laboratory and field opportunities, and for the always providing support and advice. Thank you, Dr. Craig for your support and enthusiasm regarding my research and for the opportunity to contribute to such an interesting topic in marine invertebrate ecology. I greatly appreciate this opportunity to develop my molecular techniques, my analytical skills, and my professionalism while contributing to the scientific community. Thank you to Dr. Frank Cipriano at the San Francisco State University Genomics/Transcriptomics Analysis Core for all your help, advice, and assistance. I would also like to thank Kent Susick, Ann Ho, Kyle Martin, and Danielle Perryman for all their time and assistance. Thank you all! This thesis was supported by the California State University Council on Ocean Affairs, Science and Technology (COAST) consortium, and National Science Foundation Biological-Oceanography funding (NSF award #1061695).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Watersipora Species History and Invasion

Invasive species are a threat to the biodiversity and productivity of ecosystems. This is particularly true for marine ecosystems, where the expanses of coastlines and strong ocean currents provide an avenue for the transport or migration of invasive species. The situation is complicated further by the actions of humans and our history of exploration and our development of commerce and technology. Transoceanic ship traffic has enabled organisms to colonize locations far beyond their native range (Rius *et al.* 2012). Many marine organisms are considered to be "cosmopolitan" in distribution, when in fact the events that facilitated the dispersal of these organisms occurred long before the concern over invasive species arose (Allen 1953; Geller *et al.* 2010).

The bryozoans in the genus of *Watersipora* (Neviani, 1895) are encrusting colonial organisms that have a planktonic larval stage and settle on the surfaces of boats, docks, and other floating structures in a process known as fouling. The combination of planktonic larvae, hull fouling, and a global distribution of harbors and vessel traffic provide *Watersipora* species with ample opportunity to colonize new ecosystems. The planktonic larvae of *Watersipora* species may be transported in the ballast water of ships (Carlton & Geller 1993; Drake & Lodge 2004). Colonies may also be transported by hull fouling (Carlton & Hodder 1995; Geller *et al.* 2008) which allows them to be established in locations beyond their native ecosystem's dispersal range. Carlton & Geller (1993) sampled 159 cargo ships in Coos Bay, Oregon that originated from different locations in Japan and found 28.6% of ship's ballast water contained bryozoans. Watts *et al.* (1998)

found that the common species abundance and the ability to foul had increased the geographic range for bryozoan species. *Watersipora* species also have demonstrated a tolerance for copper, which is a toxicant added to anti-fouling paint (Piola & Johnston 2006).

The genus of *Watersipora* (Neviani, 1895) belongs to the phylum Bryozoa, which is composed of encrusting colonial invertebrates. Among the bryozoans, *Watersipora* is one of the most invasive species (Mackie *et al.* 2012). The species *Watersipora subtorquata* (d'Orbigny, 1852) is of particular interest with respect to understanding the invasions of marine ecosystems worldwide. Currently *W. subtorquata* is distributed worldwide and little is known about their native range (Mackie *et al.* 2012). Previous research suggested that *W. subtorquata* is native to the Caribbean-Atlantic region (Soule & Soule 1976, cited in Mackie *et al.* 2006). In its native range, *W. subtorquata* may have been displaced by another bryozoan species, *W. subovoidea* (Mackie *et al.* 2012). Variability in environmental conditions, such as temperature and salinity, between origin and introduction locations makes marine invasions difficult to predict (Kaluza *et al.* 2010). Despite the lack of information regarding the native range and invasion timeline of *Watersipora* species, it is possible to use genetic analysis to answer questions regarding the ecology of this particular organism specifically, how it migrates through its environment, and the history of its invasion.

Analysis of the mitochondrial DNA cytochrome *c* oxidase 1 (COI) gene has revealed that *W. subtorquata* is a cryptic species complex composed of two clades of *W. subtorquata* and a cryptic unidentified species, referred to as *W.* new species (Mackie *et*

al. 2006; Láruson *et al.* 2012; Mackie *et al.* 2012). The three COI haplogroups (*W. subtorquata* clades A and B and *W.* new species) are morphologically similar and presently distinguished only by molecular analysis (Mackie *et al.* 2006; Láruson *et al.* 2012; Mackie *et al.* 2012). Analysis of COI sequences of *Watersipora* species indicated that multiple introductions from multiple source locations have occurred (Mackie *et al.* 2012). The COI haplogroup clade A has a range that includes western North America, Europe, Australasia, and Pacific Asia (Mackie *et al.* 2012). The COI haplogroup of clade B has been found in western North America and Pacific Asia (Mackie *et al.* 2012). The COI haplogroup new species has also been found in western North America and Pacific Asia (Mackie *et al.* 2012). The three COI haplogroups of the *Watersipora* species complex have been observed along the California coast in statistically distinguishable sea surface temperatures ranges (Mackie *et al.* 2012). The distribution of *Watersipora* COI haplogroups shifts from new species, clade A, and clade B, respectively along the northern coast of California where sea surface temperatures are coolest, to the southern coast where sea surface temperatures are warmer (Mackie *et al.* 2012).

Microsatellite analysis has not been used to examine population structure in *Watersipora* species along the California coast. The publication of microsatellite primers by Mackie *et al.* (2014), as part of this work, has provided the opportunity to compare microsatellite data with COI data for *Watersipora* populations. Microsatellites have been used to examine hybridization among a variety of organisms, including North American canids (*Canis lupus, C. latrans*, and *C. rufus*) (Roy *et al.* 1994), domesticated quail (*Coturnix coturnix japonica*) and wild quail (*C. c. coturnix*) (Barilani *et al.* 2005), small

Indian mongooses (*Herpestes auropunctatus*) and grey mongooses (*H. edwardsii*) (Thulin *et al.* 2006), and wild and domestic relatives of canids (*C. lupus*), felines (*Felis silvestris*), and galliformes (*Alectoris* spp.) (Randi 2008). Microsatellite analysis is also effective at detecting invasions of exotic species, similar to founder events, which are identified by a reduction in genetic diversity (Cornuet & Luikart 1996; Dlugosch & Parker 2008; Reusch *et al.* 2010). Reusch *et al.* (2010) analyzed microsatellites in populations of the comb jellyfish (*Mnemiopsis leidyi*) to identify sources of jellyfish invasion. Goldstien *et al.* (2013) combined microsatellite data with mitochondrial DNA data collected from different temporal samples to observe shifts in allele frequencies and assign new invaders to source populations.

Research Questions and Study Objectives

Coastal ecosystems are particularly sensitive to invasions of exotic species. Coastal regions surrounding locations of major shipping traffic are subject to potentially massive invasions of organisms through the movement of ships and transport of ballast water (Carlton & Geller 1993; Carlton & Hodder 1995; Drake & Lodge 2004; Geller *et al.* 2008). Carlton & Geller (1993) found samples of all marine trophic groups in their study of ship ballast water. Understanding the nature of biological invasions in ecosystems is important to preserving native environments and organisms. The anthropogenic aspect of invasions is important to determining the effect of commerce, recreation, and natural resource management on the environment.

This study used mitochondrial DNA and nuclear DNA analyses to address questions regarding the invasion history of *Watersipora* species and the structure of

populations along the California coast. "*Watersipora subtorquata"* and "*Watersipora* new species" will be referred to collectively as "*Watersipora* species" when both species are being referred to as a group. "*Watersipora* new species" will be abbreviated as "*W. n.* sp." according to Mackie *et al.* (2012). This thesis will examine using *Watersipora* populations along the California coast: 1) evidence of genetic structure and diversity, 2) genetic evidence of hybridization in the mitochondrial and nuclear data, and 3) genetic evidence of points of introduction. This study will thus increase understanding of the dynamics of the invasion and population structure of *Watersipora* species.

METHODS

Sample Collection

Watersipora species colonies were collected from eight locations along the coast of California. Sample locations include Crescent City, Humboldt Bay, Tomales Bay, San Francisco Bay, Santa Cruz, Monterey Bay, Oxnard, and San Diego Bay (Figure 1, Table 1). Samples were collected from settlement panels, docks, floats, or other structures floating in the sample locations. Sample locations were selected to provide coverage of major ports and marinas along the California coast, as well as correspond with additional *Watersipora* species research projects (data not included). Upon collection, samples were labeled, placed in 90% – 95% ethanol and transported to the laboratory to be inventoried and prepared for DNA extraction. Voucher specimens were retained for each sample collected.

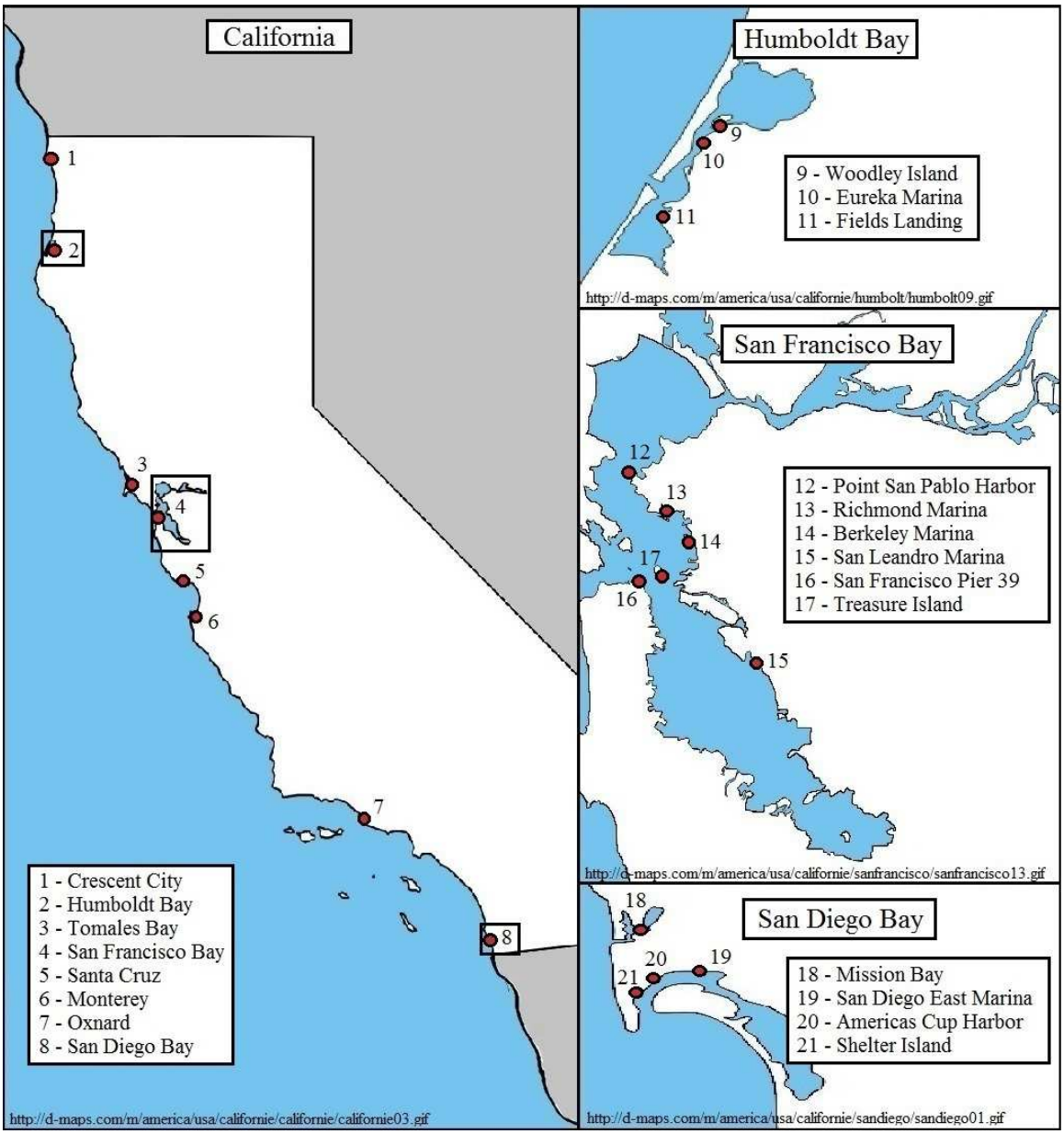

Figure 1. Map of sample collection locations. Locations where samples were collected are represented by red dots on the maps. Base maps modified with permission from d-maps.com on December 8, 2014.

Table 1. Sample collection location summary. Sample location, patch description, latitude and longitude coordinates, and number of samples collected (*n*) for each sampling location is summarized.

LOCATION	PATCH	LATITUDE	LONGITUDE	*n*
Crescent City				**27**
Crescent City	Docks	41.7459	-124.1837	27
Humboldt Bay				**48**
Woodley Island	Boat hulls	40.8074	-124.1658	5
AW2 (Woodley Island)	Docks	40.8062	-124.1621	9
AW4 (Woodley Island)	Docks	40.8061	-124.1673	7
Eureka Marina	Docks	40.8035	-124.1769	11
Fields Landing	Pilings	40.7263	-124.2215	16
Tomales Bay				**11**
Tomales Bay Resort and Marina	Docks	38.6277	-122.5144	11
San Francisco Bay				**73**
Berkeley Marina	Dock K	37.5155	-122.1850	12
Richmond Marina	Dock D	37.5445	-122.2048	6
Pier 39	Dock B	37.4832	-122.2440	6
Pier 39	Dock D	37.4835	-122.2433	8
Pier 39	Dock G	37.4841	-122.2434	6
Point San Pablo Harbor	Site 2	37.5750	-122.2508	2
Point San Pablo Harbor	Site 29	37.5746	-122.2508	4
San Leandro Marina	Dock A	37.4146	-122.1135	7
San Leandro Marina	Docks B/C/D	37.4151	-122.1128	6
San Leandro Marina	Docks G/H	37.4152	-122.1126	6
Treasure Island Sailing Center	Sailing Dock	37.4907	-122.2152	10
Santa Cruz				**36**
Santa Cruz Harbor	Dock M	36.9665	-122.0019	36
Monterey Bay				**23**
Monterey Docks	Docks	36.6039	-121.8907	23
Oxnard				**27**
Channel Islands Harbor	Docks	34.1663	-119.2268	27
San Diego Bay				**43**
San Diego East Marina	Dock H	32.7260	-117.1908	4
San Diego East Marina	Dock K	32.7260	-117.1895	5
San Diego Americas Cup Harbor	Dock 5	32.7244	-117.2249	2
San Diego Americas Cup Harbor	Site 1	32.7243	-117.2233	6
San Diego Americas Cup Harbor	Site 2	32.7248	-117.2240	6
Shelter Island (Kona Kai Marina)	Dock G	32.7133	-117.2293	10
Mission Bay	Dock	32.7677	-117.2348	10
Total				**288**

DNA Extraction

Sample DNA was extracted using a CTAB DNA extraction method (Appendix A). After extraction, the quantity and purity of DNA samples were evaluated by measuring DNA concentration (ng/µl) and DNA/protein ratio (260 nm/280 nm) using a NanoDrop ® spectrophotometer. Samples with large concentrations of DNA (> 80.0 ng/µl) were diluted to approximately 20.0 – 40.0 ng/µl for PCR amplification.

Polymerase Chain Reaction (PCR)

Amplification of microsatellites and COI sequences was performed using a standard polymerase chain reaction (PCR) protocol and visualized by agarose gel electrophoresis. PCR master mix was prepared for each locus individually (Appendix B). PCR protocols and agarose gel electrophoresis parameters were adjusted to provide the optimal amplification and visualization of PCR products. Thermal cycler amplification protocols were created for PCR amplification parameters for COI and microsatellite amplification (Appendix C). Microsatellite PCR reactions were performed individually using a combination of locus-specific primers and fluorescently-labeled locus-specific primers (Table 2). Amplification of PCR reactions was confirmed by electrophoresis and gel visualization using 2.0% agarose gels.

Table 2. Primer sequences for polymerase chain reaction (PCR) amplification. The primer sequence, sample size (*n*), and primer reference are listed for each locus. Modified primers for capillary sequencing are indicated.

Locus	Primer Sequence (5' → 3')	*n*	Primer Reference
COI	F: CATAACAGGAAGAGGTTTAAG R: TGTTGGTATAGAATAGGATC	206	Mackie *et al.*, 2006
Ws-M-1	F: TTTGAATACCTGTGTGTGTGCG (α) R: GCGTGATTGAATAAAGTTCCCC	281	Mackie *et al.*, 2014
Ws-M-2	F: GCTCTAGCTGCAATTGTCTTTCC R: TCTCCCGTACACTCTCTCTCCC (β)	284	Mackie *et al.*, 2014
Ws-M-3	F: GGATGTTGTCATTACCCTTATTGG R: TGCAGTTTGATACAATTCATCAGC (γ)	165	Mackie *et al.*, 2014
Ws-M-4	F: GGCGATTTATCGCTTCTCGG R: CGATAATTTAAAGCGCCGCC (β)	278	Mackie *et al.*, 2014
Ws-MD-1	F: GTCAGTTTTCAGACCATATT (γ) R: CACCATTATTGATCACTACA	269	Mackie *et al.*, 2014
Ws-MD-2	F: GAATCACAGTAGTTTGTCT (γ) R: TTCATATTTTAGTCATTTTATT	283	Mackie *et al.*, 2014

α: Primer labeled with HEX dye
β: Primer labeled with FAM dye
γ: Primer labeled with NED dye

DNA Sequencing

Mitochondrial DNA Sequencing

COI PCR products were purified for sequencing using EXOSAP-IT (USB Corporation) enzyme purification. COI PCR products were sent to Sequetech DNA Sequencing Services (Mountain View, CA) for sequencing. Chromatograms of sequenced PCR products were manually checked for accuracy using the software Chromas Lite v. 2.1 (Technelysium Pty. Ltd.) and verified sequences were aligned using the software MEGA v. 5.2 (Tamura *et al.* 2011).

Nuclear DNA Sequencing

Microsatellite PCR products were prepared for capillary sequencer analysis by pooling PCR products according to fluorophore dyes (Table 2). Microsatellite pooling was arranged to minimize overlap and influence of individual fluorophore dyes. Microsatellites were sequenced using an ABI 3100 capillary sequencer (Applied Biosystems) at San Francisco State University Genomics/Transcriptomics Analysis Core. PCR fragment sizes were manually checked from chromatograms using the software Peak Scanner 2 v. 2.0 (Applied Biosystems).

Genetic Analysis

COI Analysis

A set of 206 samples of *W. subtorquata* and *W. n*. sp. (*n* = 139 and 67, respectively) was sequenced, in one direction generally, from either the forward or reverse primer (Table 2). The sequences of the COI gene compared were truncated to 503 bases. When appropriate, the reverse complement of the sequenced PCR product was generated and used for sequence alignment and analysis. The truncated COI sequences were used to generate a haplotype network using the program ARLEQUIN v. 3.5.1.2 (Excoffier *et al.* 2005) and the minimum spanning tree algorithm (Rohlf 1973). ARLEQUIN was used to calculate Nei's mean number of pairwise differences, π (Nei & Li 1979) and the corrected mean number of pairwise differences between *W. subtorquata* and *W. n*. sp. COI sequences, the mean number of pairwise differences among *W. subtorquata* COI sequences, and the pairwise F_{ST} (Weir & Cockerham 1984) comparisons for *W. subtorquata* clade A COI sequences.

Microsatellite Analysis

Six microsatellite loci were selected from Mackie *et al.* (2014) for analysis (Tables 2 and 3). Deviations from Hardy Weinberg equilibrium, linkage disequilibrium, and the presence of null alleles were evaluated using GENEPOP v. 4.2 (Rousset 2008). The probability test (exact test) was used to evaluate deviations from Hardy Weinberg equilibrium with Markov Chain Monte Carlo (MCMC) parameters set to 10,000 dememorizations, 500 batches, and 10,000 iterations per batch. Linkage disequilibrium between loci was evaluated using Fisher's method between pairs of loci with MCMC parameters set to 10,000 dememorizations, 500 batches, and 10,000 iterations per batch. The presence of null alleles was estimated using a maximum likelihood method. Allele frequencies, observed and expected heterozygosity values, and polymorphic information content values (PIC) were calculated using CERVUS v. 3.0.3 (Kalinowski *et al.* 2007) (Table 3). Allele scores at each locus were compared using original locus-specific primer and the complementary fluorescently-labeled primer, and the PIGtailed primer (Brownstein *et al.* 1996) and complementary fluorescently-labeled primer (Mackie *et al.* 2014). The microsatellite data was checked for allele scoring errors by selecting a random 10% of all samples for repetition of the PCR and fragment length scoring as suggested by Dewoody *et al.* (2006). If samples initially selected by the randomization process lacked sufficient volume of DNA required for PCR amplification of all markers, replacements were selected from a randomized list of additional samples. Genotyping mismatches were reported as mismatches per allele for each locus and all loci and also

reported as mismatches per reaction for each locus and all loci (Dewoody *et al.* 2006) (Appendix D).

Table 3. Descriptive statistics for microsatellite loci. The repeat motifs, allele size (base pairs), number of samples (*n*), number of alleles (*K*), observed heterozygosity (H_O), expected heterozygosity (H_E), and polymorphic information content (PIC) are listed for each locus.

Locus	Motif	Allele Size (bp)	*n*	*K*	H_O	H_E	PIC (*W. subtorquata*)	PIC (*W. n.* sp.)	PIC (Overall)
Ws-M-1	GT	116 - 295	281	23	0.381	0.712	0.323	0.684	0.688
Ws-M-2	AG	116 - 181	284	17	0.708	0.786	0.658	0.571	0.761
Ws-M-3	TA	210 - 454	165	6	0.061	0.218	0.202	0.000	0.202
Ws-M-4	AGGT	314 - 550	278	86	0.388	0.955	0.960	0.791	0.952
Ws-MD-1	TGTC	427 - 465	269	13	0.372	0.790	0.498	0.692	0.760
Ws-MD-2	ACAT	272 - 292	283	6	0.187	0.555	0.023	0.355	0.490

The program STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000) uses Bayesian methods to calculate the proportion of the genome for each sample that originated in the assumed number of populations. STRUCTURE uses microsatellite data and an assumed number of populations, *K*, which can be estimated using a variety of methods. The value of *K* was estimated using the Δ*K* method (Evanno *et al.* 2005), which requires data from multiple runs of STRUCTURE assuming different *K* values. Each run was performed using *K* values from 1 to 10. Five iterations were performed for each *K* value with a burn-in length of 100,000 repetitions and 100,000 Markov Chain Monte Carlo (MCMC) repetitions after burn-in. Samples were analyzed in three scenarios: 1) all *Watersipora* species samples ($n = 288$), 2) all *W. subtorquata* samples ($n = 173$), and 3) all *W. n*. sp. samples ($n = 115$). The program STRUCTURE HARVESTER v. 0.3 (Earl 2012) was used to estimate *K* for each scenario using the LnP(D) L'(*K*), L"(*K*), and $|L''(K)|$ values from each iteration for each *K* value. The optimal value for *K* was visualized by plotting Δ*K* vs. *K*, where the optimal *K* value is represented by a peak in the graph.

Evidence for genetic variation among different hierarchical levels was performed using an analysis of molecular variance (AMOVA) in the program GenAlEx v. 6.5 (Peakall & Smouse 2012). AMOVA tests were completed using 9,999 permutations. Genetic structure was evaluated among regions by using regions defined by the California State Coastal Conservancy (http://scc.ca.gov) (North Coast Region, Central Coast Region, South Coast Region, and San Francisco Bay Area). The North Coast Region includes Crescent City and Humboldt Bay, the Central Coast Region includes Santa Cruz and Monterey, the South Coast Region includes Oxnard and San Diego Bay,

and the San Francisco Bay Area includes Tomales Bay and San Francisco Bay (Figure 1). Fixation indices were calculated according to Weir & Cockerham (1984) using formulas and notation by Peakall *et al.* (1995) using GenAlEx. Fixation indices were tested for significance using 9,999 permutations. By adding regional information to the AMOVA analysis, five fixation indices were calculated to investigate demographic relationships among sample populations at the levels of the individual (F_{IT}), the population (F_{ST}), the total population (F_{IS}), within regions (F_{SR}), and among regions (F_{RT}) (Peakall *et al.* 1995). The inbreeding coefficient (F_{ST}), overall inbreeding coefficient (F_{IT}), and fixation index (F_{ST}) are based on the fixation indices by Wright (1978). Tests for genetic differentiation among regions (F_{RT}) shuffled whole populations within regions, while tests for genetic differentiation among populations within regions (F_{SR}) shuffled individuals within regions (Peakall & Smouse 2012). In addition to fixation indices, the estimate of differentiation, *D* (Jost 2008) was calculated in the program GenAlEx according to Meirmans & Hedrick (2011). The statistic *D* performs better at measuring differentiation from allele frequency data compared to F_{ST} and related statistics (Jost 2008; Meirmans & Hedrick 2011). Jost's *D* is based on the effective number of alleles which has a linear relationship with diversity, as opposed to heterozygosity (Jost 2008; Meirmans & Hedrick 2011). Compared to *F*-related statistics, Jost's *D* performs better measuring differentiation when alleles become fixed (Meirmans & Hedrick 2011). However, *D* is not capable of estimating hierarchical population structure (Meirmans & Hedrick 2011).

Effective Population Size Estimation

The effective population size (N_e) was estimated for *W. subtorquata* and *W. n*. sp. using the programs LDNe v. 1.31 (Waples & Do 2008) and ONeSAMP v. 1.2 (Tallmon *et al.* 2008). LDNe uses Burrow's method for estimating linkage disequilibrium (Burrow's Δ), which requires neither assumptions about random mating nor haplotype frequencies (Waples & Do 2008). The effective population size was estimated for each *Watersipora* species using the random mating system parameter (instead of lifetime monogamy), and each species was analyzed as a single population. Three different estimates for N_e are generated based on the value used as the lowest allele frequency (0.05, 0.02, or 0.01). Alleles with frequencies lower than the value of the lowest allele frequency are excluded from the analysis. The lower the allele frequency value, the greater number of independent comparisons that can be made to estimate N_e. ONeSAMP is a web-based program that uses summary statistics and Bayesian computation to estimate effective population size from microsatellite data (Tallmon *et al.* 2008). ONeSAMP estimates N_e by creating 50,000 simulated populations, each with a randomly selected value of initial level of genetic variation (between 2 and 12) and number of generations (between 2 and 8) which reproduce according to a Wright-Fisher model (Tallmon *et al.* 2008). The effective population size is estimated from a weighted local regression of the accepted N_e values from the simulated populations that have summary statistic values close to those of the target population (Tallmon *et al.* 2008).

Clonal Analysis

Samples involve the collection of established colonies from docks and submerged structures and colonies recently settled on panels. Collection of larger, established colonies may increase the probability of sampling colonies of identical genotypes as compared to collection from settlement panels where larvae have established new colonies. The occurrence of identical genotypes was analyzed due to the sample collection methods and the asexual reproductive abilities of *Watersipora* species. The *Watersipora* species microsatellite data were analyzed for the presence of identical genotypes using GenAlEx. Only samples with complete genotypes for all six loci could be analyzed for the presence of clones, as samples with incomplete genotypes could not accurately be compared to samples with complete genotypes. Matching genotypes were assigned a letter, and sample locations of clones and genotype statistics were summarized for *W. n*. sp. and *W. subtorquata*.

RESULTS

COI Analysis

The truncated DNA sequences and translated amino acid sequences were consistent with an externally provided *W. subtorquata* voucher specimen (GenBank accession: AF441083). All of the mitochondrial DNA sequences coded for amino acids and no stop codons were observed in the sequence translation. The 503 base pair sequences varied at 28 nucleotides and only six amino acid positions differed (96.4% identical positions), considering the 206 samples analyzed. The 206 mitochondrial DNA sequences consisted of nine unique haplotypes.

The mitochondrial DNA analysis separates the DNA sequences into the three distinct mitochondrial haplogroups identified by Mackie *et al*. (2012) (Figure 2). Of the nine unique haplotypes observed, seven haplotypes belonged to the COI haplogroup *W. subtorquata* clade A, and the haplogroups *W. subtorquata* clade B and *W. n*. sp. were represented by a single haplotype (Figure 2). The *W. subtorquata* clade A haplogroup consisted of seven haplotypes, the dominant haplotype was represented by 54 samples, the second most frequently observed haplotype consisted of 49 samples, the third most frequently observed haplotype consisted of 19 samples, and the remaining four haplotypes consisted of a single sample each (Figure 2). The *W. subtorquata* clade B haplogroup consisted of 13 identical samples. The *W. n*. sp. haplogroup consisted of 67 identical samples.

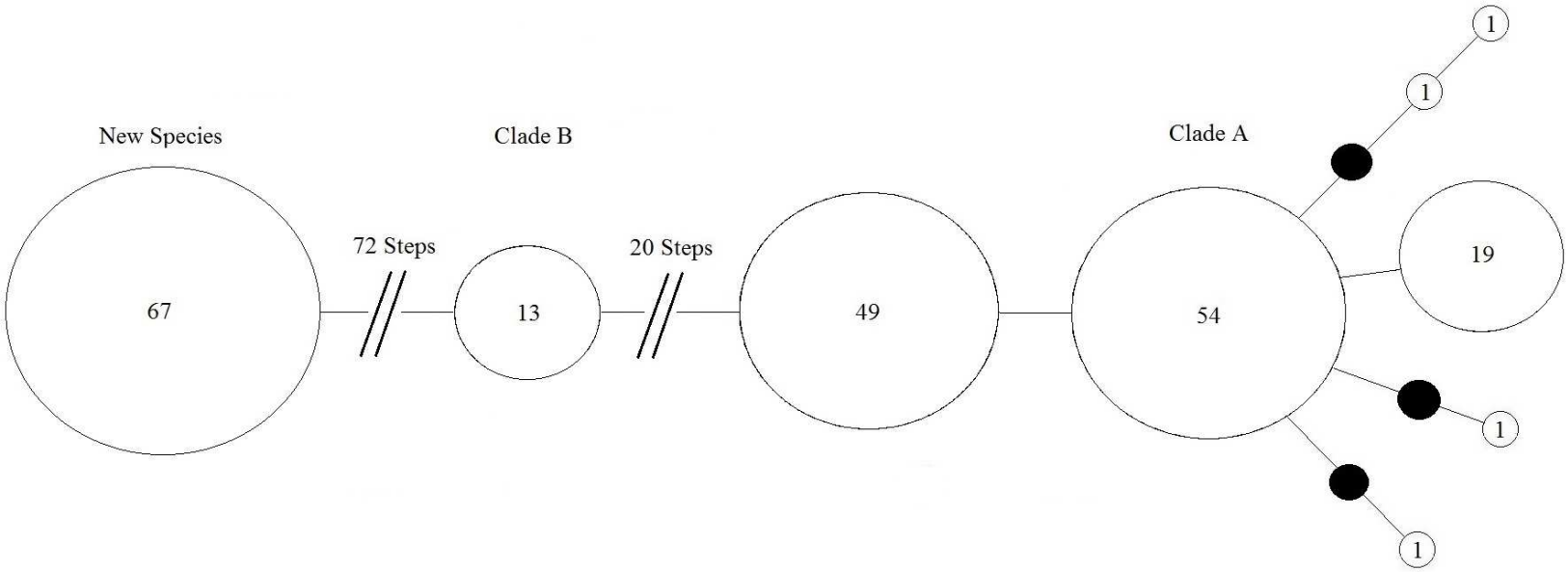

Figure 2. Haplotype network of mitochondrial DNA sequences for *Watersipora* species. *Watersipora* species COI haplogroups are labeled. Each line connecting network components represents a difference of a single nucleotide. Black circles represent non-sampled hypothetical intermediate haplotypes. Parallel lines with labels between haplotypes *W. subtorquata* clade B and *W. n.* sp., and haplotypes *W. subtorquata* clade A and *W. subtorquata* clade B haplotypes indicate a series of uninterrupted steps between haplotypes. The number of samples observed for each haplotype is indicated inside each circle.

The mean number of pairwise differences (π) and corrected mean number of pairwise differences between *W. subtorquata* and *W. n*. sp. COI sequences were 75.094 ($p = 0.00$) and 72.950 ($p = 0.00$), respectively. The mean number of pairwise differences (π) among *W. subtorquata* COI sequences was 4.286. The proportion of each haplogroup observed was plotted for each sample location (Figure 3). Pairwise F_{ST} values were calculated for *W. subtorquata* clade A and clade B mitochondrial DNA sequences among the six sampling locations where the both clades haplogroup was observed (Table 4). Pairwise F_{ST} values were not calculated for *W.* n. sp. mitochondrial DNA sequences due to lack of variation in the haplogroup. All sample locations had high average F_{ST} values, the largest average F_{ST} value was Humboldt Bay (0.364) and the smallest average F_{ST} value was Santa Cruz (0.105) (Table 4). All five pairwise F_{ST} comparisons involving Humboldt Bay or San Francisco Bay were large and statistically significant (Table 4). Negative F_{ST} values (indicating high similarities) were observed between San Diego Bay and Oxnard (-0.014) and Santa Cruz and Oxnard (-0.029) (Table 4).

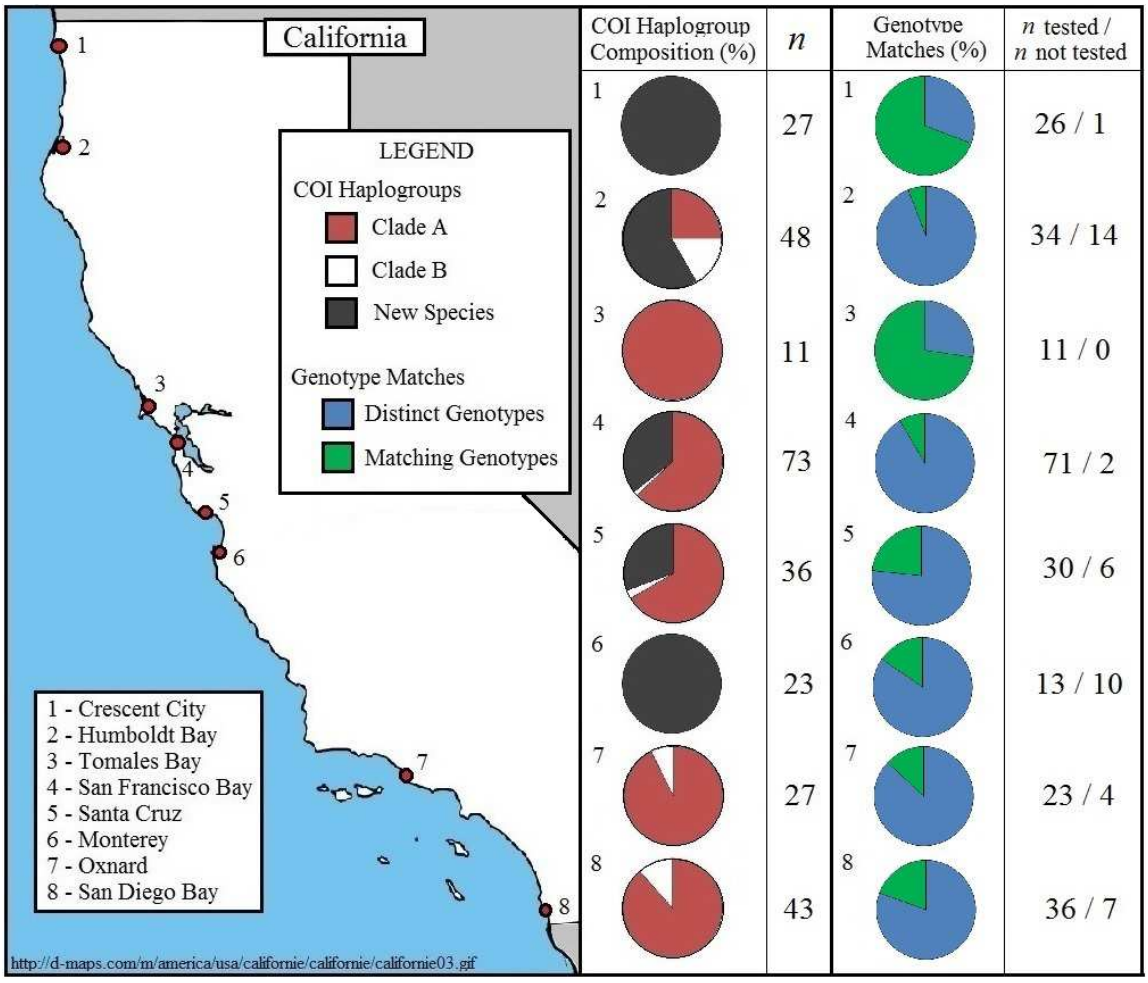

Figure 3. Map of *Watersipora* COI haplogroup distribution and matching genotypes. Sample size (COI haplogroup composition) and number of samples tested and not tested (genotype matches) are indicated next to the corresponding pie chart. Base map modified with permission from d-maps.com on December 8, 2014.

Table 4. Pairwise comparisons of F_{ST} values for *W. subtorquata* COI sequences. Pairwise F_{ST} values are below the diagonal and *p*-values are above the diagonal. Statistically significant F_{ST} values identified by * ($p < 0.05$). Calculations are based on 1,023 permutations. Average value of all pairwise F_{ST} values for each sampling site is shown in the bottom row.

	HB	TB	SF	SC	OX	SD
HB		0.006	0.000	0.000	0.001	0.010
TB	0.398*		0.000	0.406	0.176	0.063
SF	0.552*	0.580*		0.000	0.000	0.000
SC	0.346*	0.002	0.193*		0.611	0.248
OX	0.320*	0.034	0.302*	-0.029		0.635
SD	0.206*	0.057	0.239*	0.011	-0.014	
Avg F_{ST}	0.364	0.214	0.373	0.105	0.123	0.202

Microsatellite Analysis

Exact tests for deviations from Hardy-Weinberg equilibrium (HWE) varied by locus and location for both *W. subtorquata* samples and *W. n*. sp. samples. When each species was analyzed separately, results indicated both species were not in HWE. The deviations from HWE are consistent with a deficiency in heterozygosity across all loci (Table 3), and evidently due to spatial pattern. Linkage disequilibrium was not detected among *W. subtorquata* samples for any sample locations or *W. n*. sp. for any sample locations except Crescent City. The Crescent City collection location was a single, small floating dock and the samples were all established, developed colonies. Asexual reproduction of neighboring groups of colonies may explain the collection of samples with identical genotypes and observed location-level heterozygote deficiency. Linkage disequilibrium was not detected when all other populations of both species were analyzed separately, suggesting the source of linkage disequilibrium was asexual reproduction

instead of alleles physically linked on the same chromosome. Further, when the duplicate genotypes were removed from the analysis, linkage disequilibrium was not detected among the *W. n*. sp. samples at Crescent City. Null alleles were estimated to occur in both *W. subtorquata* and *W. n*. sp. data sets at various frequencies among all sample locations. The presence of null alleles may account for the observed deficiency in heterozygotes and the deviations from HWE; however, they do not appear to be a primary reason for deviations from HWE.

Microsatellite genotype scores were consistent with genotype scores obtained from quality control PCR samples used to detect allele scoring errors. Genotype mismatches observed in the QC PCR samples among microsatellite loci ranged from 0 mismatches per 28 reactions (Ws-M-3 and Ws-MD-2; 0.0%) to 3 mismatches per 28 reactions (Ws-M-1, Ws-M-2, and Ws-M-4; 10.7%), with a total of 10 mismatches per 168 reactions (6.0%) (Appendix D). Mismatches among individual alleles ranged from 0 mismatches per 6 alleles per locus (Ws-M-3 and Ws-MD-2; 0.0%) to 3 alleles with mismatches out of 86 observed alleles total (Ws-M-4; 3.5%), to 2 alleles with mismatches out of 17 observed alleles total (Ws-M-2; 11.8%), with a total of 8 alleles with mismatches out of 151 observed alleles (5.3%) (Appendix D).

The Evanno *et al.* (2005) method for estimating *K* supported *K* values of 2, 2, and 5 for all *Watersipora* species samples, *W. subtorquata* samples, and *W. n*. sp. samples, respectively (Figure 4). The *Watersipora* species microsatellite data indicates the samples consist of two distinct groups (Figure 5A). These two groups correspond to the COI lineage of the samples as belonging to either the *W. subtorquata* or *W. n*. sp.

haplogroups. The STRUCTURE population assignment for all *W. subtorquata* samples describes a mixture of population proportions among the samples (Figure 5B); however, the two groups do not correspond to the COI lineages of *W. subtorquata* clade A or clade B. The *W. n*. sp. population assignment depicts a variety of population proportions, varying also with sample locations (Figure 5C). The Humboldt Bay, San Francisco Bay, and Santa Cruz *W. subtorquata* samples (Figure 5B), and the Crescent City and Monterey *W. n.* sp. samples (Figure 5C) had complex assignment patterns in STRUCTURE plots. The Crescent City and Monterey sample assignments displayed a greater number of samples with similar population assignment proportions (Figure 5C). The Crescent City samples consisted of individuals with similar or identical genotypes, perhaps due to sampling bias or clonal patches described above.

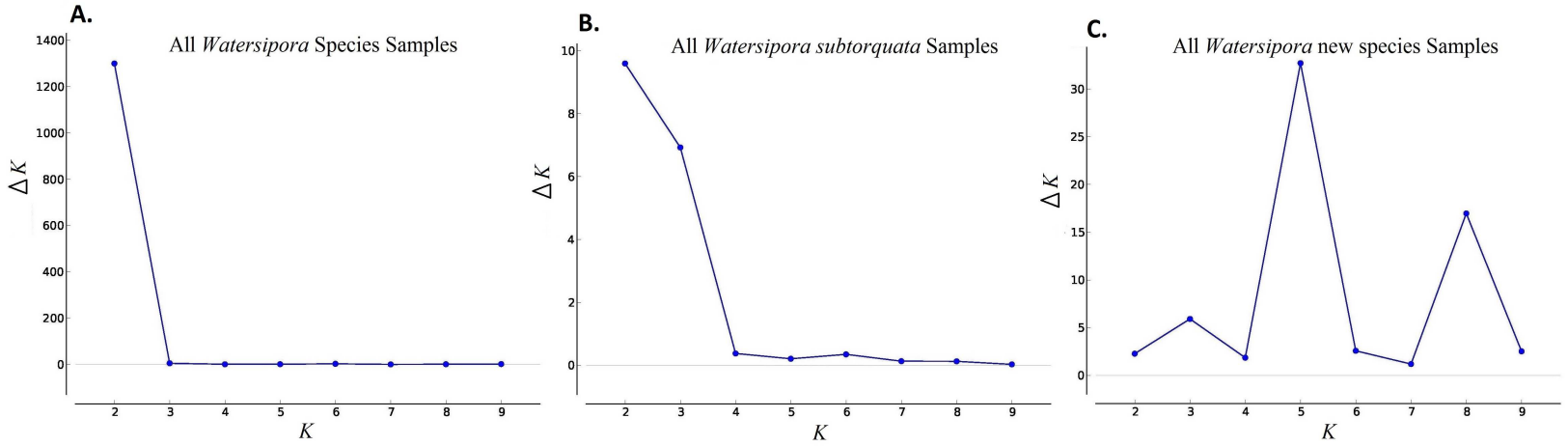

Figure 4. Estimation of *K* line graphs for *Watersipora* species. *K* estimated using the Evanno *et al.* (2005) method for three *Watersipora* species haplogroup scenarios: **A.** All *Watersipora* species samples, $K = 2$, $n = 288$; **B.** All *Watersipora subtorquata* samples, $K = 2$, $n = 173$; **C.** All *Watersipora n*. sp. samples, $K = 5$, $n = 115$.

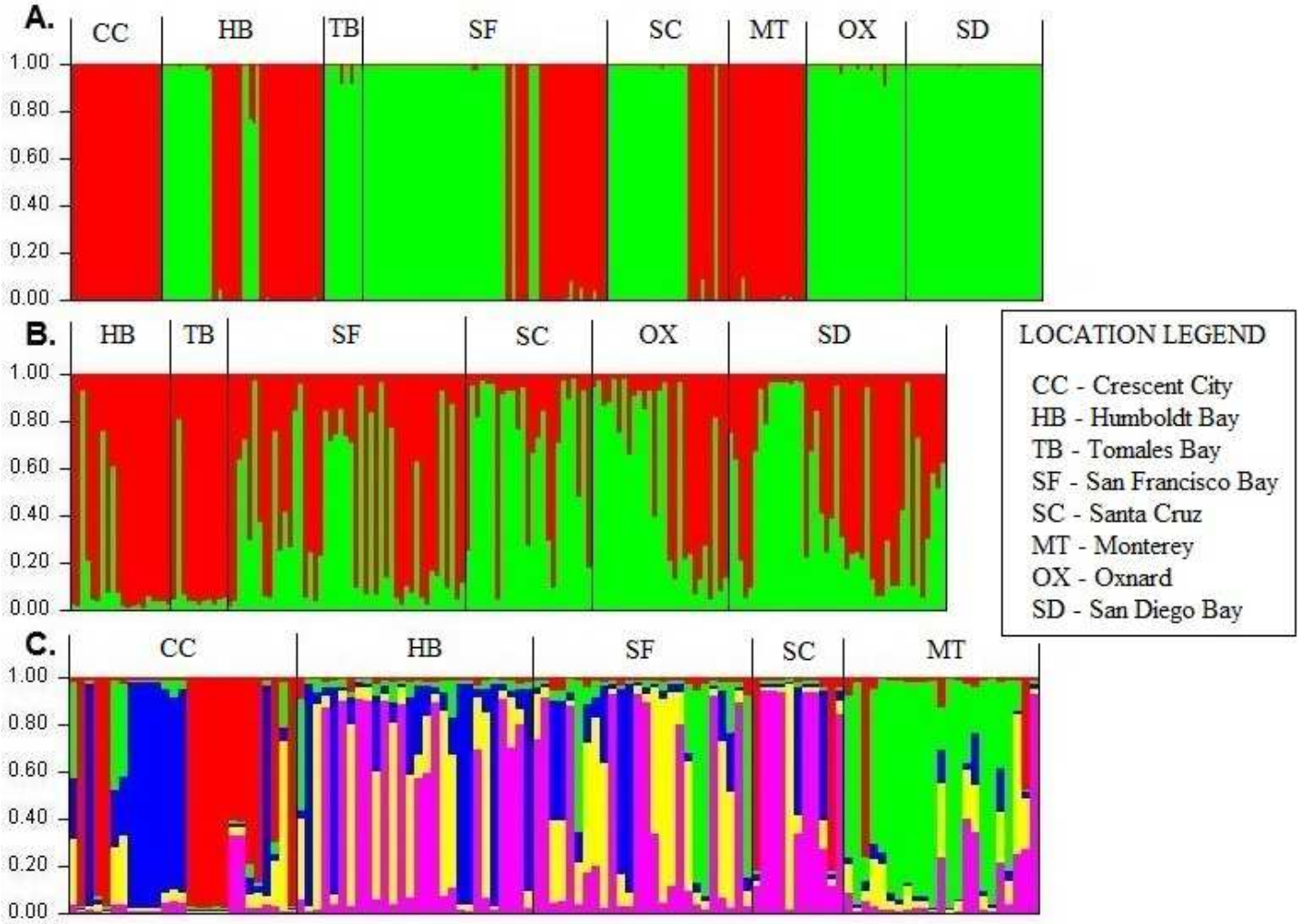

Figure 5. Population assignment bar plots for *Watersipora* species. Bayesian analysis was used for three *Watersipora* species haplogroup scenarios: **A.** All *Watersipora* species samples, $K = 2$, $n = 288$; **B.** All *Watersipora subtorquata* samples, $K = 2$, $n = 173$; **C.** All *Watersipora n*. sp. samples, $K = 5$, $n = 115$.

Tests for genetic structure were performed by AMOVA and pairwise comparisons of F_{ST} and Jost's *D* were calculated using GenAlEx. Coastal regions used to measure F_{RT} and F_{SR} are similar to those described by California State Coastal Conservancy: the North Coast Region includes Crescent City and Humboldt Bay, the Central Coast Region includes Santa Cruz and Monterey, the South Coast Region includes Oxnard and San Diego Bay, and the San Francisco Bay Area includes Tomales Bay and San Francisco Bay. Among *W. subtorquata* samples, the source of genetic variation was 4.6% among regions, 5.5% among populations, 36.2% among individuals, and 53.7% within individuals (Table 5). Among *W. n.* sp. samples, the source of genetic variation was 0.0% among regions, 10.5% among populations, 20.2% among individuals, and 69.3% within individuals (Table 5). Corresponding *F*-statistics for both species show similar trends among genetic variation at the level of individuals, populations, and regions (Table 6). The two *F*-statistics assessing genetic structure at the individual level (F_{IS} and F_{IT}) were larger than the statistics at the population (F_{ST}) and regional levels (F_{SR} and F_{RT}) in both species (Table 6). Pairwise comparisons of both F_{ST} and Jost's *D* for *W. subtorquata* were of similar size, with ranges of 0.020 to 0.234 (average = 0.118) and 0.013 to 0.230 (average = 0.117), respectively (Table 7). F_{ST} and Jost's *D* for *W. n.* sp. ranged from 0.022 to 0.140 (average = 0.085) and 0.047 to 0.250 (average = 0.159), respectively (Table 8).

Table 5. Sources of genetic variation for *Watersipora* species. Genetic variation estimated by AMOVA using GenAlEx. Calculations are based on 9,999 permutations.

Source of Variation	*W. subtorquata*		*W. n.* sp.	
	Estimated Variance	% Variance	Estimated Variance	% Variance
Among Regions	0.073	4.6%	0.000	0.0%
Among Populations	0.086	5.5%	0.192	10.5%
Among Individuals	0.570	36.2%	0.370	20.2%
Within Individuals	0.847	53.7%	1.265	69.3%
Total	1.576	100.0%	1.826	100.0%

Table 6. AMOVA *F*-statistics summary for *Watersipora* species. *F*-statistics estimated by AMOVA using GenAlEx. Test values and *p*-values are shown for each *F*-statistic. Calculations are based on 9,999 permutations.

F-Statistics	*W. subtorquata*		*W. n.* sp.	
	Value	*p*-value	Value	*p*-value
F_{RT}	0.046	0.001	-0.032	1.000
F_{SR}	0.058	0.001	0.105	0.001
F_{ST}	0.101	0.001	0.076	0.001
F_{IS}	0.402	0.001	0.226	0.001
F_{IT}	0.463	0.001	0.285	0.001

Table 7. Pairwise comparisons of F_{ST} and Jost's *D* for *W. subtorquata*. Pairwise F_{ST} values are below the diagonal and *D* values are above the diagonal. Bold values are statistically significant at the level of ($p \leq 0.001$), other statistically significant values identified by * ($p < 0.05$) or ** ($p < 0.01$). Calculations are based on 9,999 permutations. Average value of all pairwise F_{ST} and *D* values for each sampling site are shown in the bottom row.

	HB	TB	SF	SC	OX	SD
HB		**0.230**	**0.158**	**0.176**	**0.220**	**0.196**
TB	**0.221**		**0.074**	**0.152**	**0.210**	**0.117**
SF	**0.160**	**0.107**		**0.040**	**0.079**	0.013**
SC	**0.103**	**0.162**	**0.062**		0.022*	0.024**
OX	**0.161**	**0.234**	**0.093**	**0.044**		**0.044**
SD	**0.160**	**0.155**	0.020**	**0.049**	**0.043**	
Avg F_{ST}	0.161	0.176	0.088	0.084	0.115	0.085
Avg *D*	0.196	0.157	0.073	0.083	0.115	0.079

Table 8. Pairwise comparisons of F_{ST} and Jost's *D* for *W.* new species. Pairwise F_{ST} values are below the diagonal and *D* values are above the diagonal. Bold values are statistically significant at the level of ($p \leq 0.001$), other statistically significant values identified by * ($p < 0.05$) or ** ($p < 0.01$). Calculations are based on 9,999 permutations. Average value of all pairwise F_{ST} and *D* values for each sampling site are shown in the bottom row.

	CC	HB	SF	SC	MT
CC		**0.176**	**0.177**	**0.238**	**0.250**
HB	**0.095**		0.047**	**0.111**	**0.165**
SF	**0.088**	0.022*		0.089**	**0.085**
SC	**0.140**	0.067**	0.051*		**0.209**
MT	**0.123**	**0.084**	**0.055**	**0.123**	
Avg F_{ST}	0.111	0.067	0.054	0.095	0.096
Avg *D*	0.210	0.125	0.100	0.184	0.177

Effective Population Size Estimation

Estimates of effective population size (N_e) depended on the method used. Both N_e estimator programs inferred larger N_e values in *W. subtorquata*. Estimates of N_e by LDNe ranged from 63.4 to 70.5 for *W. subtorquata* and 16.0 to 37.6 for *W. n.* sp., depending on the value of the lowest allele frequency considered in the analysis (Table 9). The smallest value for the lowest allele frequency (0.01) yielded the largest number of independent comparisons (718 comparisons for *W. subtorquata* and 295 for *W. n.* sp.). Less variation was observed in N_e estimates for *W. subtorquata* and the 95% confidence intervals showed more overlap compared to *W. n.* sp. (Table 9). Estimates of N_e by ONeSAMP differed depending on the priors identified before analysis. Two estimates were completed using identical data sets and different priors. The first run used minimum and maximum N_e values of 10 and 10,000, respectively (Table 10). The second run used minimum and maximum N_e values of 4 and 500, respectively (Table 10). The estimated *W. subtorquata* N_e values for runs 1 and 2 were 3,057.04 and 7,140.78, respectively (Table 10). The estimated *W. n.* sp. N_e values for runs 1 and 2 were 194.61 and 74.84, respectively (Table 10). Estimates of N_e for each sampling location for both *Watersipora* species were calculated using LDNe (Appendix E); however, five of 11 estimates returned negative values, possibly resulting from the ratio of sample size to effective population size being too small (Brown 2012).

Table 9. LDNe effective population size estimates for *Watersipora* species. *W. subtorquata* (n = 173) and *W. n*. sp. (n = 115) calculations based on a random mating model.

Parameter	***W. subtorquata***			***W. n*. sp.**		
Lowest Allele Frequency	0.05	0.02	0.01	0.05	0.02	0.01
Independent Comparisons	96	223	718	115	205	295
Estimated N_e	63.4	70.5	68.6	16.0	23.4	37.6
Parametric 95% C.I.	(34.2, 135.2)	(45.9, 116.2)	(53.9, 89.1)	(9.8, 25.2)	(16.1, 34.2)	(26.3, 55.4)

Table 10. ONeSAMP effective population size estimates for *Watersipora* species. *W. subtorquata* (n = 167) and *W. n*. sp. (n = 100) calculations based on 50,000 iterations. Estimated N_e and associated statistics are based on two different sets of priors: † N_e estimated using priors minimum and maximum N_e = 10 and 10,000, respectively; ‡ N_e estimated using priors minimum and maximum N_e = 4 and 500, respectively.

Parameter	***W. subtorquata*†**	***W. subtorquata*‡**	***W. n*. sp.†**	***W. n*. sp.‡**
Mean Estimated N_e	3,057.04	7,140.78	194.61	74.84
Median Estimated N_e	3,712.75	7,273.21	223.10	74.01
95% C.I.	(971.88, 51,855.26)	(1,997.40, 451,329.70)	(95.47, 963.46)	(43.35, 190.85)

Clonal Analysis

Identical microsatellite genotypes were observed in both *Watersipora* species which may be consistent with clonal reproduction (Tables 11 and 12). The proportion of samples with distinct genotypes and matching genotypes was plotted for each sample location (Figure 3). Among 144 *W. subtorquata* genotypes analyzed, 24 samples matched the genotype of one or more samples (16.7%, Table 11). A single *W. subtorquata* clade B sample (CI B24R #1, Oxnard) shared a genotype with five clade A samples from San Diego Bay (Table 11). The 11 samples from Tomales Bay contained eight samples matching three different genotypes (72.7%). Among 100 *W. n*. sp. genotypes analyzed, 31 samples matched the genotype of one or more samples (31.0%, Table 12). Among the 26 Crescent City samples analyzed, 18 samples matched seven different genotypes (69.2%).

Table 11. Clonal analysis for *W. subtorquata*. Genotype labels indicate a genotype with two or more observations. The location, number of clones, and total number of clones is indicated for each genotype.

Genotype Label	Sample Location (*n*)	Total Matches
A	Tomales Bay (2)	2
B	Tomales Bay (3)	3
C	Tomales Bay (3)	3
D	San Diego Bay (5), Oxnard (1)	6
E	Oxnard (2)	2
F	San Diego Bay (1), San Francisco Bay (1)	2
G	San Francisco Bay (2)	2
H	San Francisco Bay (2)	2
I	San Diego Bay (1), San Francisco Bay (1)	2
	Total Matching Genotypes	24
	Total Unique Genotypes	120
	Total Genotypes Analyzed	144
	% Matching Genotypes	16.7

Table 12. Clonal analysis for *W.* new species. Genotype labels indicate a genotype with two or more observations. The location, number of clones, and total number of clones is indicated for each genotype.

Genotype Label	Sample Location (*n*)	Total Matches
A	Santa Cruz (2)	2
B	Crescent City (2)	2
C	Crescent City (3)	3
D	Crescent City (2)	2
E	Crescent City (3)	3
F	Crescent City (4)	4
G	Crescent City (2)	2
H	Santa Cruz (3)	3
I	Santa Cruz (2)	2
J	San Francisco Bay (2)	2
K	Monterey (2)	2
L	Humboldt Bay (2)	2
M	Crescent City (2)	2
	Total Matching Genotypes	31
	Total Unique Genotypes	69
	Total Genotypes Analyzed	100
	% Matching Genotypes	31.0

DISCUSSION

This thesis combined mitochondrial DNA analysis and nuclear DNA analysis to examine population structure, hybridization, and points of invasion among *Watersipora* species populations along the California coast. Distribution of COI haplotypes corresponded with previously reported patterns in California (Mackie *et al.* 2012). Mitochondrial and nuclear DNA analyses correctly identified samples based on COI haplogroup. Notwithstanding one highly variable locus which shared allele sizes between the *W. subtorquata* and *W. n.* sp. COI mitochondrial lineages, there was no evidence of microsatellite allele intergradation of the *W. n*. sp. and *W. subtorquata* COI defined lineages, and therefore no evidence of hybridization. Genetic variation among samples did not indicate a point of introduction; rather the microsatellite data show low hierarchical variation among sample locations and regions for each species.

Coastal Population Structure

The observed distribution of *Watersipora* species mitochondrial haplogroups was similar to the distribution patterns in previous work by Mackie *et al.* (2012). The haplogroups *W. subtorquata* clade A and clade B were observed along the California coast, from Humboldt Bay to San Diego Bay (north to south, respectively). The distribution of *W. n*. sp. ranged from Crescent City to Monterey Bay (north to south, respectively). The COI mitochondrial haplogroup clade A was the most observed often (n = 157), followed by the haplogroup *n*. sp. (n = 115), then the haplogroup clade B (n = 16). Santa Cruz was the only location represented in all nine *Watersipora* species COI

haplotypes. Pairwise comparisons of F_{ST} values calculated from *W. subtorquata* clade A COI sequences showed little genetic differentiation between Humboldt Bay and San Francisco Bay, as well as between Oxnard and San Diego Bay (Table 4). Mackie *et al.* (2012) collected *W. arcuata* in locations south of Point Conception; however this study did not observe any *W. arcuata* specimens during sample collection.

The inbreeding coefficient (F_{IS}) and overall inbreeding coefficient (F_{IT}) for *W. subtorquata* (0.402 and 0.463, respectively) and *W. n*. sp. (0.226 and 0.285, respectively) were both consistent with the deficiency of heterozygotes observed among both species within populations and among all samples (Table 6). According to general guidelines for the interpretation of F_{ST} values (Wright 1978), the fixation index for both *W. subtorquata* and *W. n*. sp. (0.101 and 0.76, respectively) indicated moderate genetic differentiation (Table 6). This interpretation of the results for F_{ST} corresponded to the AMOVA estimate of variation among populations for *W. subtorquata* and *W. n*. sp. (5.5% and 10.5%, respectively, Table 5). The genetic differentiation among populations within regions (F_{SR}) and among regions relative to the total variance (F_{RT}) for *W. subtorquata* (0.058 and 0.046, respectively) and *W. n*. sp. (0.105 and -0.032, respectively) were both consistent with low regional variance calculated by AMOVA (Table 6).

Pairwise F_{ST} comparisons indicated moderate to high genetic differentiation among all sample locations (Tables 7 and 8). Comparisons of F_{ST} among *W. subtorquata* samples showed that the Humboldt Bay and Tomales Bay samples had a greater degree of genetic differentiation relative to the other locations, whereas San Francisco Bay, Santa Cruz, Oxnard, and San Diego Bay showed a smaller degree of genetic

differentiation relative to each other (Table 7). Comparisons of F_{ST} among *W. n*. sp. samples showed the Crescent City and Monterey samples had a greater degree of genetic differentiation relative to most of the other locations, whereas Humboldt Bay, San Francisco Bay, and Santa Cruz displayed moderate genetic differentiation (Table 8). Pairwise comparisons of *D* for *W. subtorquata* and *W. n*. sp. samples showed similar trends in genetic differentiation as comparisons of F_{ST} (Tables 7 and 8). The averages of Jost's *D* among the *Watersipora* species sample locations followed the same pattern as the averages of F_{ST} (Tables 7 and 8). Among *W. subtorquata* and *W. n*. sp. average F_{ST} values, the sample location with the largest average F_{ST} value was also the sample location with the greatest observed number of clones (Tomales Bay and Crescent City, respectively, Tables 11 and 12). The relatively large number of identical genotypes may have inflated the degree of genetic differentiation at these two sample locations.

Variation was observed in the estimates for population size (N_e) for both species using different estimators (LDNe and ONeSAMP) and different priors for one estimator (ONeSAMP). Due to restrictions on samples with missing data, LDNe calculations were performed using a larger number of samples (173 *W. subtorquata* samples and 115 *W. n*. sp. samples, Table 9) than calculations performed using ONeSAMP (167 *W. subtorquata* samples and 100 *W. n*. sp. samples, Table 10). In general, the N_e estimates calculated by LDNe were smaller than the estimates calculated by ONeSAMP (Tables 9 and 10). By using a different set of priors, two prominently different N_e estimates were calculated by ONeSAMP (Table 10).

The performance of ONeSAMP is dependent upon knowledge of the populations in question, such as true N_e values or actual population sizes (Holleley *et al.* 2014). Whiteley *et al.* (2010) estimated N_e using eight microsatellite loci in populations of coastal cutthroat trout (*Oncorhynchus clarkii clarkii*) in 12 Alaskan streams containing permanent barriers to fish movement. Whiteley *et al.* (2010) used LDNe and ONeSAMP to estimate N_e values, and found the estimates from the two estimators to be significantly correlated. They used 2.0 as the minimum N_e value and both N_c (demographic estimate of population size) and 0.5 x N_c as the maximum N_e values, and found that N_e estimates were not sensitive to estimator priors (Whiteley *et al.* 2010). Holleley *et al.* (2014) suggest that ONeSAMP may provide more accurate estimates of N_e because it takes into account population information and uses eight summary statistics, where LDNe only uses Burrow's Δ. However, both estimators maybe be improved by increasing sample size and the number of loci analyzed (Holleley *et al.* 2014). Despite using only Burrow's Δ to estimate N_e, LDNe may provide a more accurate estimate of N_e because it is not biased by unknown priors. It is also possible that the sample sizes or number of loci analyzed in this study were too small to detect large effective population size values.

Patterns were observed in the settlement of larvae (or absence) for *Watersipora* species. Settlement panels deployed to observe larvae settlement among arrays containing three levels of copper-treated anti-fouling paint panels (control, low dosage, and high dosage) resulted in the settlement of only *W. subtorquata* larvae, despite deployment in regions where both *Watersipora* species were observed using other collection methods. Additional observations using arrays of settlement panels revealed

identical results in a sampling region (Crescent City) where this study observed only *W. n*. sp. by manual removal of samples from docks and similar structures (Mackie, personal communication). The observation of only *W. subtorquata* larvae settlement on panels deployed in regions where both species are observed shows the importance of understanding larval behavior and reproductive timing differences of the two species. The settlement panel design used in this study may not be capable of sampling *W. n*. sp. larvae settlement. One possibility is that larvae of *W. n.* sp. are settled on or close to the parent colony. Further investigation into the reproductive cycles of these two species may provide insight to the observed absence of *W. n*. sp. larvae on settlement panels and new techniques to sample larvae of both species. This study illustrates the importance of understanding reproductive differences at the parental and larval level for purposes of invasive species management and control of invasive species dispersal.

Hybridization Assessment

Three regions were sampled where all three COI haplogroups coexist along the California coast (Humboldt Bay, San Francisco Bay, and Santa Cruz; Figure 1), where the probability of observing hybrids is greatest among the sample locations. Among the samples collected in San Francisco Bay, *W. n*. sp. colonies were only collected from locations near the mouth of the bay. All the samples collected at Pier 39 and 60.0 % of the samples from Treasure Island Sailing Center were *W. n*. sp., however *W. n*. sp. accounted for only 35.6% of the samples collected in San Francisco Bay. The samples collected at Santa Cruz and Humboldt Bay were 30.6% and 58.3% *W. n*. sp., respectively.

W. subtorquata and *W. n*. sp. samples were more intermixed at these two locations, the most intimate example of this being a colony of *W. subtorquata* clade A growing on a colony of *W. n*. sp. collected from Santa Cruz (SC M1A and SC M1B, respectively). After DNA extraction and COI and microsatellite PCR, these two samples have COI haplotypes and microsatellite genotypes typical of the appropriate species and share no alleles in common. When applied to the entire *Watersipora* species dataset, population assignments made by STRUCTURE correctly separated the entire dataset into groups consisting of *W. subtorquata* and *W. n*. sp. with few samples containing a partial population assignment (22 samples, 7.64%, Figure 5A). Among these 22 samples, only two samples had a partial population assignment 80% or less (Humboldt Bay, Hum Eu 788 and Hum Eu 790). Both of these samples had missing data for one locus (Ws-MD-1), otherwise the genotypes for these two samples were typical of *W. subtorquata* samples in that region. All the samples observed with partial population assignments were either missing data for one or more loci, or was affected by one highly variable locus (Ws-M-4). Some samples possessed incomplete genotypes due to depletion of DNA samples. Marker Ws-M-4 was highly variable among *Watersipora* species with a total of 86 observed alleles. By comparison, the marker with the second largest number of alleles was Ws-M-1, with a total of 23 observed alleles, or approximately one-third as many alleles (Table 3). Observed genotypes of the six loci were generally typical of either *W. subtorquata* or *W. n*. sp., despite one highly variable locus and rare alleles.

The *Watersipora* species nuclear and mitochondrial datasets were congruent. Mitochondrial clade typing PCR (Láruson *et al.* 2012) and COI gene sequencing were

consistent in all cases (when acceptable PCR amplification and sequencing results were achieved). Three of the microsatellite markers tested (Ws-M-2, Ws-MD-1, and Ws-MD-2) were capable of accurately identifying sample COI haplogroup individually.

Case studies using similar techniques and of similar scope using mitochondrial DNA and nuclear (microsatellite) DNA data include the stalked sea squirt (*Styela clava*) (Goldstien *et al.* 2013), domesticated quail (*Coturnix coturnix japonica*) and wild quail (*C. c. coturnix*) (Barilani *et al.* 2005), and North American wild canids (*Canis* spp.) (Roy *et al.* 1994). Similarly to *Watersipora* species, the ascidian *S. clava* is transported globally via hull fouling (Goldstien *et al.* 2013). Goldstien *et al.* (2013) showed consistency between mitochondrial and nuclear DNA analysis when assigning individual samples from recent invasions to populations previously sampled. Their data showed also that timing of sampling can record the appearance and disappearance of alleles from a population (Goldstien *et al.* 2013). Barilani *et al.* (2005) observed Japanese quail (*C. c. japonica*) mitochondrial haplotypes and an admixture of microsatellite alleles in Common quail (*C. c. coturnix*) populations. Their research reports the evidence of mitochondrial and nuclear DNA in viable hybrids (Barilani *et al.* 2005). Previous analysis of mitochondrial DNA (Lehman *et al.* 1991) and microsatellites (Roy *et al.* 1994) in North American wild canids has revealed evidence of hybridization among gray wolves (*C. lupus*) and coyotes (*C. latrans*). Lehman *et al.* (1991) found coyote mitochondrial genotypes in populations of gray wolves, but gray wolf mitochondrial genotypes were absent from coyote populations. Roy *et al.* (1994) found that genetic distance between all coyote populations and hybridizing gray wolf populations were

significantly less that all coyote populations and nonhybridizing gray wolf populations. The proportion of coyote alleles shared with gray wolves was 76 of 92 alleles (82.6%) and the proportion of gray wolf alleles shared with coyotes was 76 of 95 alleles (80.0%) (Roy *et al.* 1994).

The microsatellite data from this thesis did not contain any shared allele combinations of markers with alleles typical of each species. The absence of samples from our dataset with a particular mitochondrial lineage with a nuclear genotype containing at least one marker typical of a different mitochondrial lineage (for example, a sample with a *W. subtorquata* mitochondrial lineage with a microsatellite genotype containing at least one locus typical of *W. n*. sp.) suggests that if these two species do hybridize, hybrids are rare or perhaps not viable.

Invasion Assessment

The locations sampled in this study represent a large distance of coastline greater than the distance traversable by a marine organism with planktonic larvae. The potential for hull fouling and rafting (Carlton & Hodder 1995), for larvae to be transported by ballast water (Carlton & Geller 1993; Drake & Lodge 2004), or the tolerance for copper in anti-fouling paints (Piola & Johnston 2006) have provided opportunities for bryozoans, such as *Watersipora* species, to expand their distribution beyond their native range (Watts *et al.* 1998). An introduction of an invasive species at a given location (a type of founder effect), and the absence of gene flow or additional introductions would potentially create a population of unique allele frequencies (Cornuet & Luikart 1996;

Dlugosch & Parker 2008; Reusch *et al.* 2010). If a successful introduction of an invasive species migrated or expanded beyond the original point of introduction, according to founder-effect principles, the greatest genetic diversity (or number of alleles) would occur at the original point of introduction, and genetic diversity would decrease as the range of the invasive species expanded.

Among *W. subtorquata* samples, the number of observed alleles are (from largest to smallest), San Francisco Bay, 60 alleles; San Diego Bay, 52 alleles; Humboldt Bay, 44 alleles; Oxnard, 43 alleles; Santa Cruz, 33 alleles; and Tomales Bay, 14 alleles. Among *W. n*. sp. samples, the number of observed alleles are (from largest to smallest), San Francisco Bay and Humboldt Bay, 31 alleles; Monterey, 22 alleles; Crescent City, 20 alleles, and Santa Cruz, 18 alleles. The San Francisco Bay populations for both species display the greatest number of alleles, and Humboldt Bay populations display the second largest number of alleles for *W.* new species and the third largest number of alleles for *W. subtorquata*, the second largest number of alleles for *W. subtorquata* are observed in San Diego Bay. San Francisco Bay and San Diego Bay are both large hubs for commercial and private boat traffic, acting as ports for international shipping traffic and as local marinas that serve private and commercial vessels. Considering the activity of commercial and private boat traffic along the California coast since the first observation of *Watersipora* species in 1980s, the potential for subsequent introductions from these two main ports seems extremely high. If colonies or larvae of *Watersipora* species have been consistently introduced along the Pacific Coast since the 1980s, it seems plausible

that anthropogenic transportation and introductions have created an effect similar to gene flow that has masked the unique genetic signature created by any original introductions.

AMOVA results indicated that the greatest source of variation is among individuals, instead of the levels of populations or regions (Table 5). This may be a result of the continuous addition of genetic material from the transportation of colonies or larvae to these regions from beyond their natural range of dispersal. This effect may also be observed in the low value of *K* calculated from the search for optimal *K* values for analysis in STRUCTURE.

The STRUCTURE analysis of *W. subtorquata* samples indicates an inferred population of two, and the population assignment scores do not suggest any significant structure among sample locations (Figures 4 and 5). The STRUCTURE analysis of *W. n*. sp. samples indicates an inferred population of five, and the population assignment scores display a greater mixture of inferred populations in the Humboldt Bay, San Francisco Bay, and Santa Cruz locations compared to the Crescent City and Monterey locations, which each display a large proportion of a couple or single dominant inferred populations. These patterns tentatively suggest that San Francisco Bay and Humboldt Bay may be points of introduction for *W. n.* sp. and San Francisco Bay, San Diego Bay, and Humboldt Bay may be points of introduction for *W. subtorquata*.

The oceanic currents off the California coast create biogeographic boundaries for many marine taxa (Burton & Lee 1994; Dawson 2001; Wares *et al.* 2001; Kelly & Palumbi 2010; Mackie *et al.* 2012). The California coast ocean currents consist of the California Current, a southward flow of cold, low salinity, highly oxygenated water; and

the Southern California Counter Current or Davidson Current, a northward flow of warm, high-salinity, low-oxygen water (Dawson 2001). These two currents interact near Point Conception, California (34.5˚ N) where geographic and environmental conditions create a region of hydrographic complexity (Dawson 2001; Wares *et al.* 2001). Results from separate investigations differ on the significance of Point Conception as a boundary for marine taxa. Dawson (2001) states that Point Conception is not a major biogeographic boundary and that the Monterey Bay region and the Los Angeles region represent phylogenetic breaks along the California coast. Wares *et al.* (2001) describes Point Conception as a "leaky" boundary for northern species with pelagic larval dispersal and states that Point Conception acts as a boundary for southern taxa more so than northern taxa. Seasonal relaxation of upwelling and El Niño events may increase the movement of pelagic larvae northward past Point Conception, while the California Current may carry larvae southward past Point Conception (Kelly & Palumbi 2010). Previous research describes the Point Conception region to act as a phylogeographic break for populations of *Watersipora* species (Mackie *et al.* 2012) and the marine copepod *Tigriopus californicus* (Burton & Lee 1994). Kelly & Palumbi (2010) found four out of 41 species with pelagic larvae sampled to show genetic differentiation across Point Conception.

The AMOVA results for both *Watersipora* species do not show strong genetic differentiation at the regional or population level (Tables 5 and 6). Pairwise comparisons of F_{ST} and Jost's D do not show trends of genetic differentiation increasing with distance between sampling locations, or between sampling locations separated by Point Conception (Tables 7 and 8). The AMOVA analysis, indicating and an overall lack of

genetic differentiation among sample regions, support the occurrence of frequent transportation of *Watersipora* larvae and colonies by commercial and private vessel traffic along the coast and beyond.

CONCLUSION

The distribution of COI haplogroups observed in this study is consistent with that reported by Mackie *et al.* (2012). This study did not observe the *W. arcuata* COI haplogroup in southern California reported by Mackie *et al.* (2012); however, the distributions of the mitochondrial haplogroups *W. subtorquata* clade A and clade B, and *W. n.* sp. were observed in similar ranges. Sea surface temperature data were not measured as a part of this thesis, but COI haplogroup observations from this study are consistent with those from Mackie *et al.* (2012). Mitochondrial haplogroup distributions from this study support sea surface temperatures as a factor limiting distributions of *Watersipora* mitochondrial haplogroups.

The microsatellite data in this study did not clearly identify exact locations of introduction for *W. subtorquata* or *W. n.* sp.; however, the results of this study suggest the greatest source of genetic variation is within individuals for both species. This level of genetic variation may support scenario of the regular transportation of genetic material by commercial and private vessel traffic along the California coast. Nuclear and mitochondrial analyses in this study did not find any evidence of hybridization between these two species of *Watersipora*, such has *W. subtorquata* mitochondrial lineages with *W. n.* sp. alleles or genotypes, and vice versa. More research is needed to determine the

barriers to hybridization between these two species. A greater understanding of temporal and environmental conditions for spawning and the behavior of larvae of both species, as well as studies of chromosome numbers and gamete compatibility may be relevant in addressing the potential for hybridization between the two species.

REFERENCES

Allen F (1953) Distribution of marine invertebrates by ships. *Marine and Freshwater Research*, **4**, 307-316.

Barilani M, Deregnaucourt S, Gallego S, Galli L, Mucci N, Piombo R, Puigcerver M, Rimondi S, Rodríguez-Teijeiro J, Spanò S (2005) Detecting hybridization in wild (*Coturnix c. coturnix*) and domesticated (*Coturnix c. japonica*) quail populations. *Biological Conservation*, **126**, 445-455.

Brown A (2012) Estimating the Effective Population Size of *Crassostrea virginica*. Cornell University, doctoral dissertation.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques*, **20**, 1004-1006, 1008-1010.

Burton RS, Lee B-N (1994) Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. *Proceedings of the National Academy of Sciences*, **91**, 5197-5201.

Carlton J, Hodder J (1995) Biogeography and dispersal of coastal marine organisms: experimental studies on a replica of a 16th-century sailing vessel. *Marine Biology*, **121**, 721-730.

Carlton JT, Geller JB (1993) Ecological Roulette: The Global Transport of Nonindigenous Marine Organisms. *Science*, **261**, 78-82.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Dawson MN (2001) Phylogeography in coastal marine animals: a solution from California? *Journal of Biogeography*, **28**, 723-736.

Dewoody J, Nason JD, Hipkins VD (2006) Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology Notes*, **6**, 951-957.

Dlugosch K, Parker I (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431-449.

Drake JM, Lodge DM (2004) Global hot spots of biological invasions: Evaluating options for ballast–water management. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, 575-580.

Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47.

Geller J, Mackie J, Schroeder G, Gerhinger D (2008) Distribution of highly invasive bryozoans belonging to a cryptic species complex in the genus *Watersipora* determined by DNA sequences. *Final Report to California Department of Fish and Game, Moss Landing Marine Laboratories, CA*.

Geller JB, Darling JA, Carlton JT (2010) Genetic perspectives on marine biological invasions. *Annual Review of Marine Science*, **2**, 367-393.

Goldstien SJ, Inglis GJ, Schiel DR, Gemmell NJ (2013) Using temporal sampling to improve attribution of source populations for invasive species. *PloS One*, **8**, e65656.

Holleley CE, Nichols RA, Whitehead MR, Adamack AT, Gunn MR, Sherwin WB (2014) Testing single-sample estimators of effective population size in genetically structured populations. *Conservation Genetics*, **15**, 23-35.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099-1106.

Kaluza P, Kölzsch A, Gastner MT, Blasius B (2010) The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, **7**, 1093-1103.

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PloS One*, **5**, e8594.

Láruson ÁJ, Craig SF, Messer KJ, Mackie JA (2012) Rapid and reliable inference of mitochondrial phylogroups among *Watersipora* species, an invasive group of ship-fouling species (Bryozoa, Cheilostomata). *Conservation Genetics Resources*, **4**, 617-619.

Lehman N, Eisenhawer A, Hansen K, Mech LD, Peterson RO, Gogan PJ, Wayne RK (1991) Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution*, **45**, 104-119.

Mackie JA, Darling JA, Geller JB (2012) Ecology of cryptic invasions: latitudinal segregation among *Watersipora* (Bryozoa) species. *Scientific Reports*, **2**, 871, DOI: 810.1038/srep00871.

Mackie JA, Keough MJ, Christidis L (2006) Invasion patterns inferred from cytochrome oxidase I sequences in three bryozoans, *Bugula neritina*, *Watersipora subtorquata*, and *Watersipora arcuata*. *Marine Biology*, **149**, 285-295.

Mackie JA, Wostenberg D, Doan M, Craig SF, Darling JA (2014) High-throughput Illumina sequencing and microsatellite design in *Watersipora* (Bryozoa), a complex of invasive species. *Conservation Genetics Resources*, **6**, 1053-1055.

Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. *Molecular Ecology Resources*, **11**, 5-18.

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, **76**, 5269-5273.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, **28**, 2537-2539.

Peakall R, Smouse PE, Huff D (1995) Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloe dactyloides*. *Molecular Ecology*, **4**, 135-148.

Piola RF, Johnston EL (2006) Differential resistance to extended copper exposure in four introduced bryozoans. *Marine Ecology Progress Series*, **311**, 103-114.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology*, **17**, 285-293.

Reusch TB, Bolte S, Sparwel M, Moss AG, Javidpour J (2010) Microsatellites reveal origin and genetic diversity of Eurasian invasions by one of the world's most notorious marine invader, *Mnemiopsis leidyi* (Ctenophora). *Molecular Ecology*, **19**, 2690-2699.

Rius M, Turon X, Ordóñez V, Pascual M (2012) Tracking invasion histories in the sea: facing complex scenarios using multilocus data. *PloS One*, **7**, e35815.

Rohlf FJ (1973) Hierarchical clustering using minimum spanning tree. *Computer Journal*, **16**, 93-95.

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553-570.

Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) Onesamp: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, **8**, 299-301.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731-2739.

Thulin C-G, Simberloff D, Barun A, Mccracken G, Pascal M, Islam MA (2006) Genetic divergence in the small Indian mongoose (*Herpestes auropunctatus*), a widely distributed invasive species. *Molecular Ecology*, **15**, 3947-3956.

Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753-756.

Wares JP, Gaines S, Cunningham CW (2001) A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution*, **55**, 295-306.

Watts PC, Thorpe JP, Taylor PD (1998) Natural and anthropogenic dispersal mechanisms in the marine environment: a study using cheilostome Bryozoa. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **353**, 453-464.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Whiteley AR, Hastings K, Wenburg JK, Frissell CA, Martin JC, Allendorf FW (2010) Genetic variation and effective population size in isolated populations of coastal cutthroat trout. *Conservation Genetics*, **11**, 1929-1943.

Wright S (1978) Evolution and the genetics of populations. Variability within and among natural populations. Vol. 4. *University of Chicago press, Chicago, IL, USA*.

Appendix A - CTAB DNA Extraction Method

1. Add 80 µl 2-mercaptoethanol to 40 ml CTAB buffer (see recipe below) right before use (0.2%) and mix.

2. Homogenize the tissue sample with microtube/blue pestle, or dice with razor blade. (Alternatively, Using liquid N_2, freeze dry sample and grind to dust.)

3. Obtain the final tissue sample in 700 µL of CTAB isolation buffer. All steps can be carried out in a 1.5 mL test tube. Incubate (60 ˚C, 1 hour minimum). (If including Pro-K use 500 µg/mL end concentration, i.e. c. 18 µL from 20 mg/mL stock tubes).

4. Add 700 µL Chloroform/Iso-amyl-alcohol (25:1 mixture, premixed reagent). Invert to mix. Open bottle in fume hood.

5. Spin 10 min at maximum speed (14000 x g) at 4°C.

6. Gently remove upper, clear aqueous phase with a P1000, or P-smaller pipet. Be careful not to disturb the interface. Discard the remaining fraction in organic waste receptacle. (Add 1 µL RNase (DNase-free) - though not really necessary if PCR and sequencing steps follow)

7. Note volume of recovered aqueous phase. Add 0.08 volumes cold 7.5 M ammonium acetate.

8. Add 0.54 volumes of 7 isopropanol: 1 ammonium acetate (2/3 of the recovered volume) to DNA to precipitate. Invert 20-30 times. Incubate on ice for 30-40 minutes. (make up 40 mL: 5 mL ammonium acetate, 35 mL isopropanol).

9. Centrifuge for 3 minutes.

10. Discard supernatant into isopropanol chemical waste jar. Be careful not to dislodge pellet.

11. Add 700 µL 70% EtOH, invert tubes 5-10 times.

12. Centrifuge for 1 minute.

13. Discard supernatant; be careful not to dislodge pellet.

14. Add 700 µL 95% EtOH, invert tubes 5-10 times.
15. Centrifuge for 1 minute.

16. Discard supernatant; be careful not to dislodge pellet.

17. Dry pellet (important to remove all ethanol prior to PCR). I leave tubes open to allow evaporation of residual ethanol. Placing the tubes on a 50˚C heat block for 3 minutes may be helpful, but don't bake DNA. Or invert tubes on a clean Kimwipe and allow to dry for 10-15 minutes upside down, or until pellet looks dry. If the pellet dried too long upside down, it will fall out. Continue to dry upright but covered by a Kimwipe for 30-45 minutes.

18. Hydrate pellets with 30-50 µL TE buffer for longer storage, or H_2O. Note, Tris is a PCR inhibitor, so when diluting in TE it is often helpful to dilute the starting sample prior to PCR. Allow DNA to resuspend overnight at room temperature. Store the DNA in the refrigerator the next day, and in freezer (-20˚C) over longer term.

19. Optional - check DNA on 1% Agarose gel – should be a high molecular weight band, indicating unsheared DNA; quantify using NanoDrop®.

Appendix B - Polymerase Chain Reaction (PCR) Master Mix Formulas

Microsatellite PCR Master Mix Formula

Reagent	**Volume (µl)**
GoTaq Green Master Mix	10.0
Nuclease-free H_2O	3.6
$MgCl_2$	2.0
BSA	0.4
Primer 1	1.0
Primer 2	1.0
Total Volume	18.0

COI PCR Master Mix Formula

Reagent	**Volume (µl)**
GoTaq Green Master Mix	10.0
Nuclease-free H_2O	4.0
$MgCl_2$	1.6
BSA	0.4
Primer 1	1.0
Primer 2	1.0
Total Volume	18.0

COI Clade Typing Master Mix Formula

Reagent	**Volume (µl)**
GoTaq Green Master Mix	5.0
Nuclease-free H_2O	3.0
Typing Primer Mix	1.0
Total Volume	8.0

Appendix C - Polymerase Chain Reaction (PCR) Amplification Protocols

I. Microsatellite 55°C Annealing Temperature Amplification Protocol:

Cycle #	Repetition	Temp.	Time
Cycle 1	1x	95°C	4:00
Cycle 2	30x	95°C	0:30
		55°C	0:30
		72°C	0:40
Cycle 3	1x	72°C	7:00
Cycle 4	1x	12°C	holding time

II. Microsatellite 50°C Annealing Temperature Amplification Protocol:

Cycle #	Repetition	Temp.	Time
Cycle 1	1x	95°C	4:00
Cycle 2	30x	95°C	0:30
		50°C	0:30
		72°C	0:40
Cycle 3	1x	72°C	7:00
Cycle 4	1x	12°C	holding time

III. COI Amplification Protocol:

Cycle #	Repetition	Temp.	Time
Cycle 1	1x	95°C	4:00
Cycle 2	10x	95°C	0:30
		55°C	0:30
		72°C	0:40
Cycle 3	25x	95°C	0:30
		45°C	0:30
		72°C	0:40
Cycle 4	1x	72°C	5:00
Cycle 5	1x	12°C	holding time

IV. COI Clade Typing Protocol:

Cycle #	Repetition	Temp.	Time
Cycle 1	1x	95°C	4:00
Cycle 2	30x	95°C	0:30
		52°C	0:30
		72°C	0:40
Cycle 3	1x	72°C	7:00
Cycle 4	1x	12°C	holding time

Appendix D - Quality Control PCR Genotype Scoring Mismatches

QC PCR Parameter	Microsatellite Loci						Total
	Ws-M-1	**Ws-M-2**	**Ws-M-3**	**Ws-M-4**	**Ws-MD-1**	**Ws-MD-2**	
# of Alleles (*K*)	23	17	6	86	13	6	151
# of Reactions	28	28	28	28	28	28	168
# of Mismatches	3	3	0	3	1	0	10
# of Alleles With Mismatches	2	2	0	3	1	0	8
Mismatches/Allele	0.087	0.118	0.000	0.035	0.077	0.000	0.053
Mismatches/Reaction	0.107	0.107	0.000	0.107	0.036	0.000	0.060
Mismatches - *W. subtorquata*	2	0	0	2	1	0	5
Mismatches - *W. n.* sp.	1	3	0	1	0	0	5

Appendix E - LDNe N_e Estimates by Sampling Location

Location	Effective Population Size (N_e)			
	W. subtorquata		*W. n.* sp.	
	N_e	95% C.I.	N_e	95% C.I.
Crescent City	-	-	1.4	(0.9, 2.2)
Humboldt Bay	187.5	(22.7, ∞)	-64.9	(111.8, ∞)
Tomales Bay	-6.4	(26.1, ∞)	-	-
San Francisco Bay	124	(58.5, 1,668.0)	-287.1	(37.6, ∞)
Santa Cruz	-41.9	(110.0, ∞)	1.8	(0.9, 8.5)
Monterey	-	-	12.3	(2.9, 427.0)
Oxnard	-150.9	(52.7, ∞)	-	-
San Diego Bay	59.2	(32.4, 171.7)	-	-