CRYPTIC SPECIATION AND POPULATION CONNECTIVITY IN THE PEANUT WORM *PHASCOLOSOMA AGASSIZII* (SIPUNCULA: PHASCOLOSOMATIDAE)

A Thesis

by

NATHAN DUNN JOHNSON

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Anja Schulze
Committee Members,	Jaime Alvarado-Bremer
	Gilbert T. Rowe
Interdisciplinary Faculty Chair,	Gilbert T. Rowe

December 2014

Major Subject: Marine Biology

ABSTRACT

The prevalence of undetected cryptic species in marine environments is one of the greatest obstacles to obtaining an accurate estimate of biodiversity. Morphological identification can underestimate true levels of diversity, as it does not allow for the detection of cryptic species. Sipunculans, commonly called peanut worms, are thought to contain high levels of cryptic diversity due to their conserved morphology and their paucity of taxonomically informative characters. In this thesis, we use genetic-identification techniques to examine diversity within the Pacific sipunculan *Phascolosoma agassizii*. Mitochondrial DNA sequence data shows that *P. agassizii* is comprised of two cryptic species, one isolated to the eastern Pacific and one to the western Pacific. These clades exhibit large amounts of genetic divergence and are not recovered as sister taxa, suggesting an early speciation event within the *Phascolosoma* genus. Based on the location of the original holotype, the eastern Pacific clade represents the true *P. agassizii*, whereas the western Pacific clade is an undescribed species of *Phascolosoma*.

The long-lived larval stage within *Phascolosoma agassizii*, known as a pelagosphera larva, is thought to engage in long-distance planktonic dispersal. Using a genetic technique called ISSR-PCR, we amplified non-coding polymorphic regions of the nuclear genomes of western and eastern *P. agassizii* to test for levels of population connectivity along a coastline. Our ISSR results gave evidence for three to five genetically distinct populations within the western Pacific and two genetically distinct

populations within the eastern Pacific. This data suggests that these coastal regions consist of genetically structured populations restricting intraspecies gene flow. Since the prevailing currents in these regions should connect most populations, larval behavior resulting in local recruitment may be the mechanism behind such population structure.

This research suggests that *P. agassizii* is not a single cohesive species, and in fact is comprised of two divergent clades. Each clade is isolated to a separate Pacific coast, with multiple genetically-distinct populations along each coast. This research highlights the need for combining genetic identification methods with morphological identification, especially in sipunculan taxonomy.

DEDICATION

To Mom and Dad

ACKNOWLEDGEMENTS

I would like to thank Dr. Anja Schulze for all her support and guidance throughout the past two and half years. I am incredibly grateful for the opportunity she provided me to pursue a career in science, and so very proud of the work we have accomplished. She is an exceptional advisor, a patient teacher, and one of the most intelligent people I know.

I would also like to thank the other members of my thesis committee, Dr. Gilbert Rowe and Dr. Jaime Alvarado-Bremer for their valuable insight and assistance throughout my research. To be mentored by three experts in their respective fields is a luxury that I was lucky enough to have experienced.

I am extremely grateful for all the support and friendship of my fellow lab mates, both past and present. Thank you to Heidi, Justin, Joseph, Liz, Russell, Marissa N., Laura, Corey, Gaby, Ren, Victoria, Marissa O., Erin, and Hannah for making the invertebrate lab not only a home for insightful systematic research but also a fun and welcoming place to work.

The images of *Phascolosoma agassizii* in this thesis were taken with cameras from the labs of Dr. Anna Armitage and Dr. Chris Marshall here at Texas A&M University at Galveston. I am thankful to both of them, as well as to Meagan, Katie, Erin, and Kelly for their assistance with the equipment.

Chapter II of this thesis is an extension of an initial study by Schulze et al. in 2012. I want to thank all those involved in this study because without their preliminary data, Chapter II would not have been possible.

This thesis represents in part a collaboration between Texas A&M University at Galveston, the Far East Branch of the Russian Academy of Sciences, the California Academy of Sciences and the Los Angeles County Museum of Natural History. I am thankful to the researchers and the staff from these institutions for all of their assistance. I particularly want to thank Anastassya Maiorova, Christina Piotrowski, and Leslie Harris for collecting and loaning us specimens for use in our genetic analyses. This work is supported by a FEBRAS-CRDF collaborative grant to Anja Schulze and Anastassya Maiorova (RUB1-2996VL-11).

I am so very grateful for all the assistance provided by those in the MARB department, especially Sarah Wall, Rachel Ball, and Christina Irons. I also want to thank Holly Richards and Nicole Kinslow for all their assistance throughout this process.

I want to extend a special note of gratitude to Dr. Mary Rice, not only for providing close to fifty years of foundational research on sipunculans, but also for sharing with me her wisdom and passion for science during my time at Friday Harbor Marine Labs.

Finally I want to thank my family, who has been incredibly supportive of my research. Aunt Marge and Uncle Jack, thank you for creating a home away from home for me in Texas. I'm sure I've racked up a dinner tab well into the triple digits, and I promise to reimburse you once my Nobel Prize money kicks in. Mom and Dad, thank

you for empowering me to pursue my interests in science. You two are a constant source of inspiration and wisdom in my personal and professional life. I'll never cease to be amazed at your level of generosity, and I love you both very much.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The concept of cryptic speciation, first introduced by Mayr (Mayr, 1948) challenges traditional morphology-based species concepts. Cryptic species are those that exhibit a high degree of genetic divergence without the presence of any distinguishing morphological characters (Knowlton, 1993). Occasionally, morphological distinctions are discovered between groups that were thought to be cryptic species. In these instances, they are referred to as pseudo-cryptic species (Knowlton, 1993). Some authors regard sibling species as synonymous with cryptic species (Knowlton, 1993; Sáez & Lozano, 2005). However in some instances cryptic lineages are not each other's most closely-related species, and therefore do not represent sibling or sister species (Knowlton, 1986; Schulze *et al.*, 2012). The phenomenon known as cryptic speciation is thought to have occurred many times in divergent evolutionary lineages, with a recent study suggesting there are anywhere between 8,000 and 35,000 instances of cryptic speciation in oceanic habitats, conservatively (Appeltans *et al.*, 2012). A review of cryptic speciation in polychaetes concluded that there were no common trends in either life history traits, morphological complexity, or ecology of cryptic species that could be used to predict the presence of cryptic diversity (Nygren, 2013).

Characterizing the actual number of species is crucial for conserving biodiversity, and the presence of cryptic species may result in an underestimation of actual biodiversity. Without a proper understanding of species ranges and demography,

we cannot effectively manage threatened species or populations. However, with the widespread adoption of rapid and inexpensive genetic identification methods based on single-gene sequencing, research facilities now have the ability to delineate taxonomic classifications previously complicated by cryptic speciation. Detection of distinct genetic lineages will not only allow for more accurate classification, but it also provides insight into the processes and mechanisms of speciation (Borda *et al.*, 2013; Nygren, 2013).

Sipunculan worms, commonly known as peanut worms, are a taxon of approximately 150 currently recognized species (Cutler, 1994). Formerly classified as their own phylum, sipunculans are now generally included as a clade within Annelida (Bleidorn, 2007; Boore & Staton, 2002; Weigert *et al.*, 2014). They are exclusively marine, benthic, and can be locally abundant, although usually in cryptic habitats (i.e. burrowed in the substrate). Peanut worms exhibit an unsegmented, vermiform body shape comprised of a trunk and an extendable introvert (Figure 1). At the anterior-most portion of the introvert lies the mouth and feeding tentacles, which surround either the mouth or the nuchal organ, a putative chemosensory structure. The anus usually lies at the anterior of the trunk but in some species the anus is located higher on the introvert (Cutler, 1994). The introvert is retracted by one to four retractor muscles, found within the coelom and attached to the anterior end of the introvert and the inner lining of the trunk (Rice, 1993). The body wall musculature consists of an outer circular layer and an inner longitudinal layer. The longitudinal musculature can either form a smooth sheath or, more frequently, is arranged into bands. The gut is typically coiled, extending from

the introvert to the trunk and then looping back around to the anus in a j-shaped fashion (Cutler, 1994).

Species delimitation using morphological characters is especially difficult for sipunculans, as their simple body plan contains few diagnostic features. Appeltans et al. (2012) estimate that 55% of the total diversity within this group may be cryptic. Stephen and Edmonds (1972) originally listed over 300 species within Sipuncula, however the currently accepted taxonomy put forth by Cutler in 1994 synonymized many of these. There are now 149 formally described sipunculan species, which were identified solely by their anatomical characters. Thus, sipunculans represent an exceptional opportunity to study genetic diversity, specifically between geographically isolated regions that may contain cryptic species. Evidence of morphospecies complexes with significant genetic divergence already exists for several species within Sipuncula (Kawauchi & Giribet, 2010; Kawauchi & Giribet, 2013; Schulze *et al.*, 2012; Staton & Rice, 1999). Additionally, much is known about the differing developmental modes of sipunculan larvae (A. Adrianov & Maiorova, 2010; Rice, 1976; Rice *et al.*, 2012; Rice & Schulze, 2004). Depending on the species, sipunculans exhibit one of four modes of larval development (Figure 2) (Schulze and Rice, 2009). These range from direct development to a two-phase larval period that includes a form capable of feeding in the water column for a few months (A. Adrianov & Maiorova, 2010; Rice, 1976). Therefore, a more accurate picture of genetic diversity within this group may reveal otherwise obscure relationships between larval developmental modes and geographic partitioning of cryptic adult forms.

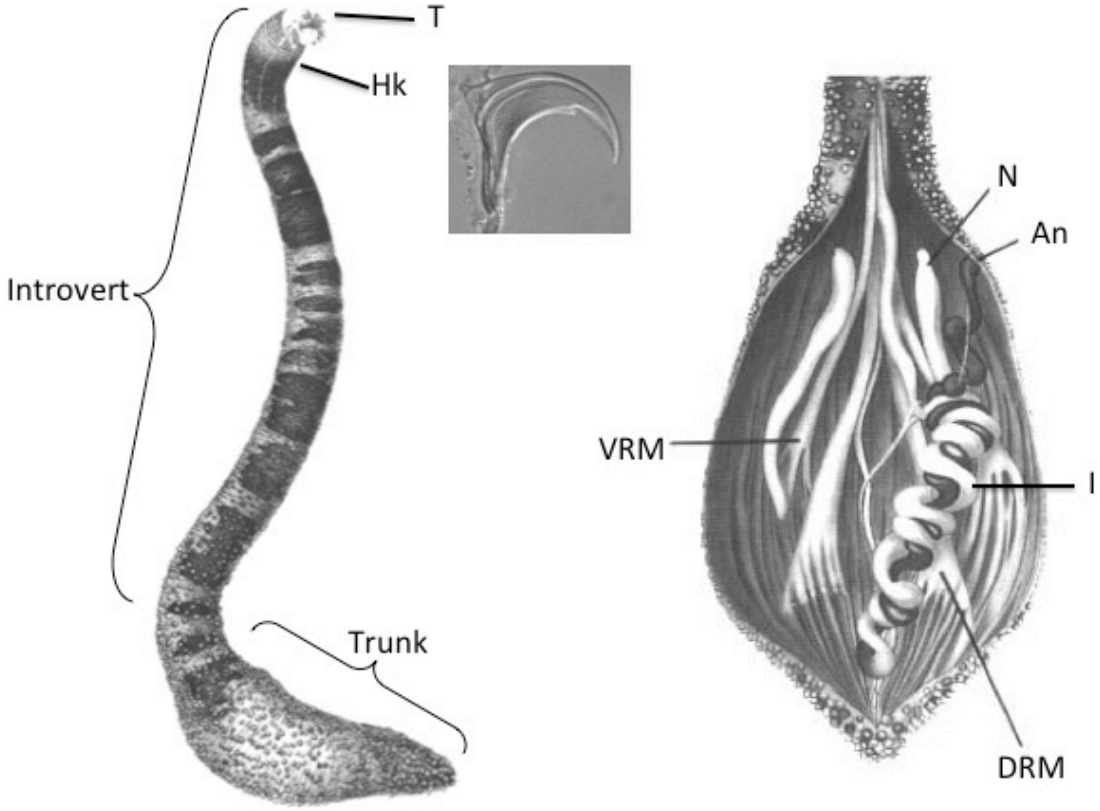

Figure 1. Morphological depiction of a dissected *Phascolosoma sp*., revealing its external and internal anatomy. T = tentacles, Hk = introvert hooks, N = nephridium, An = anus, DRM = dorsal retractor muscle, VRM = ventral retractor muscle, I = intestine. (Rice, 1993).

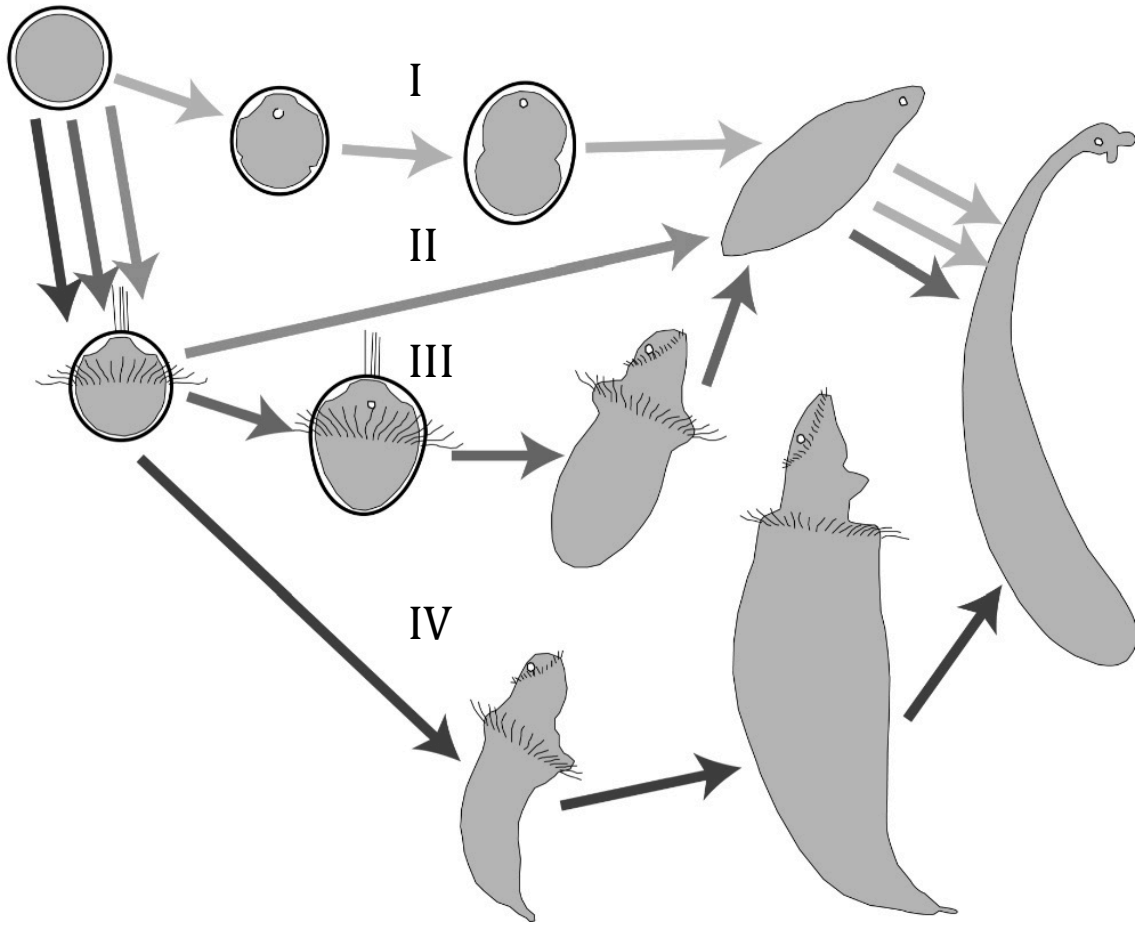

Figure 2. The four larval modes found within Sipuncula. I – direct development; II – single larval phase, a short-lived lecithotrophic trochophore; III – two larval phases, a short-lived lecithotrophic trochophore and a short-lived lecithotrophic pelagosphera; IV – two larval phases, a short-lived lecithotrophic trochophore and a long-lived planktotrophic pelagosphera. Form IV represents the development mode found within *P. agassizii*. (Schulze and Rice, 2009).

CHAPTER II

CRYPTIC SPECIATION IN *PHASCOLOSOMA AGASSIZII* (SIPUNCULA: PHASCOLOSOMATIDAE): EAST-WEST DIVERGENCE BETWEEN NON-SISTER TAXA

Overview

The use of DNA sequence data in taxonomy over the past few decades has led to the frequent detection of cryptic species in a wide variety of taxa, increasing dramatically our estimates of species diversity (Bickford *et al.*, 2007; Knowlton, 2000; Nygren, 2013). Cryptic species are morphologically similar or identical, and thus are typically classified as a single species, yet exhibit significant genetic differentiation (Nygren, 2013). Often they are sibling species. Occasionally, upon a reevaluation of their anatomy distinguishing morphological features have been identified. In these instances, they are referred to as pseudo-cryptic or pseudo-sibling species (Knowlton, 1993). Much of our current taxonomy has been based on morphological identification methods, and therefore does not reflect the presence or overall frequency of cryptic biological diversity. Not only does this lead to an inaccurate estimate of global biodiversity, but it also may affect our knowledge of a given species' distribution, reproduction strategy, and environmental tolerance, as certain cryptic species have been shown to differ in these respects (Kawauchi & Giribet, 2010; Lewis & Karageorgopoulos, 2008; Maltagliati *et al.*, 2001). The importance of utilizing integrative taxonomic approaches has been understood for quite some time (Mayr,

1948). However, it is only recently that the ability to examine an organism's genetic structure has become cost-effective (Nygren, 2013). Accordingly, identification techniques combining genetic and morphological data have been incorporated more frequently into biological research (Dayrat, 2005; Padial *et al.*, 2010). This two-pronged identification approach allows us to not only detect cryptic species, but also to re-evaluate the morphology of these organisms for the presence of any taxonomically-informative characters that previously went unnoticed (Sáez & Lozano, 2005). Further, viewing these approaches as complementary rather than mutually exclusive will promote the development of an integrative taxonomic framework that fosters a greater understanding of earth's biodiversity (Schlick-Steiner *et al.*, 2010; Will *et al.*, 2005).

Marine ecosystems are some of the most poorly sampled biomes on earth, and thus are thought to contain a large amount of cryptic diversity (Appeltans *et al.*, 2012; Bickford *et al.*, 2007). The relative scarcity of physical barriers to gene flow in the ocean compared with terrestrial environments has led to an overestimation of the geographical distribution of many marine taxa, since the potential for long-distance dispersal does not always result in actual long-distance dispersal (Knowlton & Keller, 1986). Thus, some marine species classified according to their morphology were erroneously thought to have a cosmopolitan distribution (Kawauchi & Giribet, 2013; Schulze *et al.*, 2012). Phylogeographic associations of such taxa indicate cryptic species complexes rather than a cosmopolitan distribution of a single lineage (Hoare *et al.*, 2001; Lazoski *et al.*, 2001; Solé-Cava *et al.*, 1991). In addition to being poorly sampled, marine ecosystems are also currently being threatened by climate change and

anthropogenic stressors. These variables can potentially result in decreased oceanic biodiversity and shifts in marine species' distributions (Provan & Maggs, 2012; Thomas *et al.*, 2004), which could further impact our understanding of marine taxonomy.

Introduction

Sipunculans are unsegmented marine annelids distributed throughout a wide range of depths within tropical, temperate, and polar waters (Cutler, 1994; Stephen & Edmonds, 1972). They can reach densities of approximately 300-400 individuals per square meter in certain regions (Romero-Wetzel, 1987; Thompson, 1980), and thus are thought to play a vital role in marine trophic systems. They are also active burrowers, and are believed to significantly influence the geochemistry of their respective benthic habitats (Shields & Kedra, 2009). Despite being critical components of marine ecosystems, only a few invertebrate researchers have studied some species in depth. Cutler (1994) revised sipunculan taxonomy based on morphology and reduced the number of species from over 300 (according to Stephens & Edmonds 1972) to 149. Recent studies have suggested that this may underestimate the true diversity, as genetic data has shown evidence of cryptic speciation within nominal species of sipunculans (Kawauchi & Giribet, 2010; Schulze *et al.*, 2012). Thus, there are still large gaps in our knowledge regarding the ecology, evolutionary history and intraspecies population connectivity of peanut worms. The sipunculan *Phascolosoma agassizii agassizii* (Keferstein, 1866) has been widely reported from shallow-water habitats throughout the North Pacific (Cutler, 1994). Records of this species exist from California to Alaska on the east Pacific and stretching northward from Japan to the Kamchnka Peninsula on the

west (Maiorova, pers. obs.). Additionally, there are a few records of this species along the south and west African coast in the Atlantic, as well as certain Indian Ocean localities (Cutler, 1994).

P. agassizii agassizii (Keferstein, 1866), of the family Phascolosomatidae is a benthic detritovore. It is distinguished from the subspecies *Phascolosoma agassizii kurilense* by the lack of a secondary lobe on the nephridia of mature worms (Cutler, 1994); however since there is no literature recognizing *kurilense* as a distinct subspecies since Cutler's monograph, we remain skeptical of its status as a unique subspecies. For the purposes of this study, we will refer to all specimens as simply *P. agassizii* to avoid any potential confusion. The life cycle of this species includes two pelagic larval stages prior to settling on the substrate (the figure as seen on page 5). The first is a short-lived lecithotrophic trochophore stage, while the second stage, known as pelagosphera, can spend up to several months in the water column (A. Adrianov & Maiorova, 2010; Rice, 1976). This sequence of planktonic larval stages within *P. agassizii* can potentially facilitate dispersal over long distances, and thus theoretically facilitate long distance gene flow between populations. However, an analysis of *P. agassizii's* phylogeography within the Pacific Ocean by Schulze et al. (2012) shows evidence of potential cryptic or pseudo-cryptic speciation between the east and west coasts, despite its previously believed amphi-Pacific distribution.

Differences in larval development have been observed between populations from the eastern and western Pacific. Egg sizes of *P. agassizii* in the San Juan Islands are approximately 1.5 times larger than those in the Sea of Japan. Additionally, the

spawning season and the duration of the trochophore stage are longer in eastern Pacific populations compared to those from the Sea of Japan (A. Adrianov & Maiorova, 2010). Previous research suggests this could be a result of differences in water temperatures between the two environments, since the Sea of Japan typically experiences more variability in temperature throughout the year than do the San Juan Islands (A. Adrianov & Maiorova, 2010). However, if cryptic speciation exists within this nominal species, then these differences in life history may not simply be a case of developmental plasticity.

In this study, we utilized DNA sequence data from two mitochondrial markers and one nuclear marker to examine the genetic divergence between eastern and western Pacific populations of *P. agassizii* (Figure 3).

Methods

Specimen Collection

Samples were collected from a total of nine different populations on both the eastern and western Pacific coasts (the table seen on page 16). Troitsy Bay, Amursky Bay, Ussuriysky Bay, Vostok Bay, Sokcho Bay in South Korea and Iturup Island in the Kuril Islands comprised the western Pacific sample set. Friday Harbor (WA), Whiffen Spit (BC), and California populations comprised the eastern Pacific sample set. Samples from Peter the Great Bay, Friday Harbor, and Whiffen Spit were from previous collections by Anja Schulze and colleagues, and maintained at Texas A&M Galveston. Christina Piotrowski at the California Academy of Sciences provided specimens from California. Anastassya Maiorova at the Zhirmunsky Institute of Marine Biology, Far

East Branch of the Russian Academy of Sciences provided specimens from Sokcho Bay and the Kuril Islands. Individuals were either hand-collected by SCUBA divers or collected via bottom trawling by research vessels. Except for the California specimens, which were collected at depths of 86 – 122 meters, all individuals were collected from the shallow intertidal or subtidal regions of their associated locales. All were fixed in 95% ethanol and kept at -20° C.

DNA Extraction and Sequencing

DNA was extracted from the retractor muscles of individuals according to the Qiagen DNeasy Blood and Tissue Kit® protocol. If a specimen was too small (body length < 10 mm) to obtain a sample of the retractor muscle, DNA was extracted from the body wall. 1 µl template DNA (0.22 pMol – 0.35 pMol original concentration) was combined with Gotaq® Green Master Mix (Promega) (12.5 µl), $MgCl_2$ (2.5 mM final concentration), 5' and 3' primers (0.4 mM of each), in a total volume of 25 µl to create a thermocycling master mix. For amplification of the cytochrome *c* oxidase subunit I gene, we used the forward primer LCO1490 (5' – GGTCAACAAATCATAAAGATATTG – 3') and reverse primer HCO2198c (5' – TGATTTTTTGGTCACCCTGAAGTTTA – 3') (Folmer *et al.*, 1994).

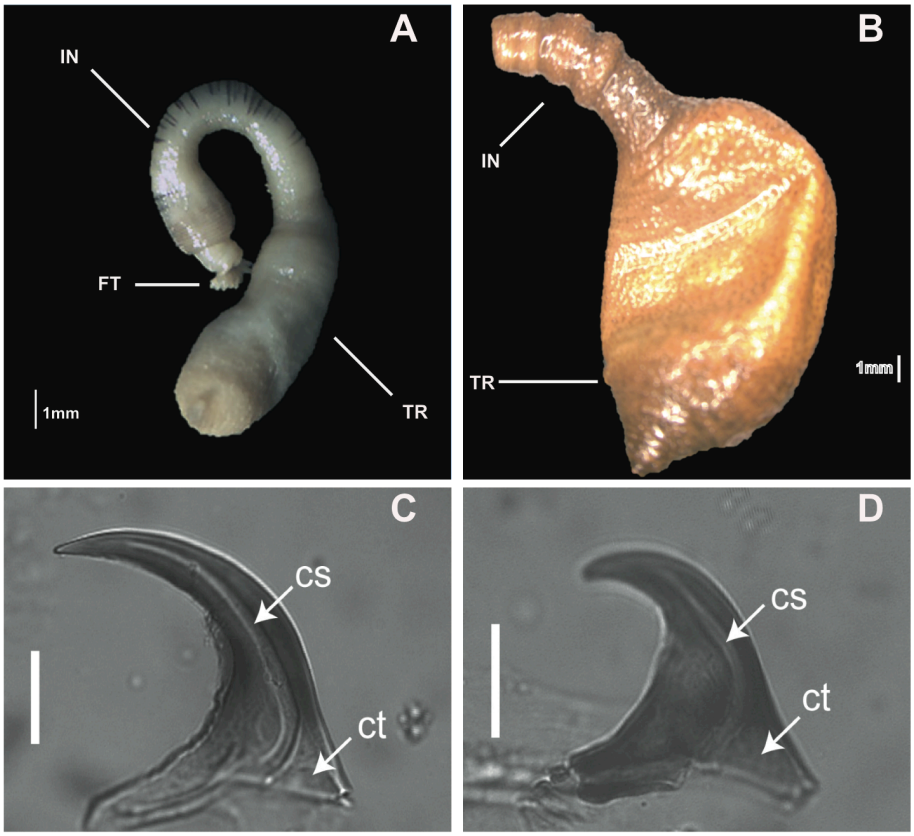

Figure 3. Images of *Phascolosoma agassizii* specimens and introvert hooks. (A) Specimen from California. Image is at 6.5X magnification. (B) Specimen from Vostok Bay in the Sea of Japan. Image is at 7.5X magnification. TR = Trunk, IN = Introvert, FT = Feeding Tentacles. Scale bars are 1 mm. (C) Introvert hooks from an eastern Pacific specimen and (D) from the western Pacific specimen. For (C) and (D), scale bars are 20 µm. CS = clear streak, CT = clear triangle. (Schulze et al., 2012).

For amplification of the 16S ribosomal RNA gene, we used the forward primer 16SPagF (5' – GCTAAGGTAGCGCAATCACT – 3') and the reverse primer 16SPagR (5' – GGGTTAGAGTGCTGCTTCAT – 3'). These 16S primers were specifically designed for *P. agassizii* over the course of this study. For amplification of the nuclear ITS gene,

we used the forward primer ITS18SFPoly (5' – GAGGAAGTAAAAGTCGTAACA – 3') and the reverse primer ITS5.8SRPoly (5' – GTTCAATGTGTCCTGCAATTC – 3') (Nygren *et al.*, 2009). The DNA was amplified under the following thermocycling protocol for the two mitochondrial genes: initial denaturing step of 94°C for 2 min, 35 cycles of denaturation at 94°C for 30 s, annealing at 40°C for 30 s, and extension at 72°C for 45 s, and a final extension at 72°C for 7 min. An annealing temperature of 45°C was used for some samples that exhibited multiple bands per lane when amplified at 40°C. The following protocol was used for amplification of the ITS gene: initial denaturing step of 94°C for 4 min, 30 cycles of denaturation at 94°C for 30 s, annealing at 45°C for 30 s, and extension at 68°C for 45 s, and a final extension at 72°C for 5 min. Following amplification, DNA products were visualized on a 1% agarose gel and cleaned with ExoSAP-IT® (Affymetrix). Cycle sequencing was performed using BigDye Terminator (Life Technologies) in 10 μL volumes and purified according to standard protocols using the BigDye X Terminator purification kit. Sequences were obtained using the ABI3130 Genetic Analyzer from Applied Biosystems, and edited using Sequencer™ v. 4.8. Editing consisted of assembling the forward and reverse strands of each specimen and removing the primer regions. All sequences were submitted to GenBank under accession numbers KM226349 – KM226482 (Table A1).

Phylogenetic Reconstruction

Sequences were aligned in MEGA 5.05 (Tamura *et al.*, 2011) using the default settings for ClustalW (Larkin *et al.*, 2007). Phylogenetic analyses were performed using

Bayesian Inference (BI), Maximum Likelihood (ML), and Neighbor Joining (NJ). Outgroups representing other *Phascolosoma* species were selected as recommended in previous studies (Schulze *et al.*, 2007; Schulze *et al.*, 2012). This was done to provide comparative sequence data for any observed levels of intraspecies divergence within our sample set, as well as to root our phylogenetic trees. *Phascolosoma granulatum* was selected as the root for our trees based on results from earlier studies examining the phylogenetic relationships of Sipuncula (Schulze *et al.*, 2005, 2007). Tree reconstruction according to Bayesian Inference was conducted in MrBayes 3.1.2 (Ronquist *et al.*, 2012) using a generalized-time reversible model incorporating a gamma distribution of rate substitution with invariant sites (GTR+I+G). The GTR+I+G model was selected as it resulted in the lowest AIC and BIC scores, calculated in jModelTest 2.4.1 (Posada, 2008). Bayesian trees were constructed over 10,000,000 generations with a 25% burn in (2,500,000 generations). Analyses using Maximum Likelihood (ML) were performed in RAxML (Stamatakis, 2014) using a GTR+I+G model. Branch support values were obtained over 1,000 bootstrap replicates. Analyses using NJ were constructed in MEGA 5.05 using the Tamura-Nei model (1993) with a gamma distribution and 1,000 bootstrap replicates.

Genetic Distance and Diversity Calculations

Average between-group genetic distances were calculated in MEGA using Kimura's 2-Parameter model (K2P) (1980) as well as the Tamura-Nei model, which was selected using jModelTest. The K2P model was selected to provide a comparison with similar studies of *P. agassizii* (Schulze *et al.*, 2012). Standard error values for these

distances were calculated using 1,000 bootstrap replicates. Estimates of interpopulation and intrapopulation diversity, as well as the coefficient of genetic differentiation (G_{ST}) were also calculated in MEGA using both the Tamura-Nei model and Kimura's 2-Parameter model. Rate variation between sites incorporated a gamma distribution model and took into account only sites that had 95% site coverage or greater. An analysis of molecular variance was performed using Arlequin v. 3.5 (Excoffier & Lischer, 2010) to compare variance between groups, between populations, and within a population. The hierarchical arrangement consisted of two groups corresponding to the eastern Pacific (East) and western Pacific (West). Collecting localities shown in Figure 4 represented populations for this analysis. Computations were run using the Tamura-Nei model (Tamura & Nei, 1993) and incorporated 1,000 permutations. The number of haplotypes and a calculation of F_{ST} were also determined in Arlequin.

Results

Alignments

Our COI dataset resulted in a 662 base-pair alignment after deletion of primer regions and gap editing. A total of 68 individuals were represented within this alignment, five outgroups within the *Phascolosom*a genus and 63 nominal *P. agassizii* specimens from nine Pacific populations (Table 1). Our 16S dataset was comprised of 58 total individuals with the same five outgroups, resulting in a 509 bp alignment. A third dataset consisting of a 373 bp alignment of the ITS-1 nuclear region for 18 individuals, representing all populations, was included in our study to explore the

possibility of obtaining different phylogenies with nDNA compared to mitochondrial DNA.

Table 1. *Phascolosoma agassizii* genetic dataset showing collection location, sample sizes for each gene sequenced, and depth.

Population	Coordinates	COI	16S	ITS	Concatenated Phylogeny	Concatenated AMOVA	Ocean	Depth (m)
Troitsy Bay (TB)	42.64°N 131.04°E	10	10	2	10	10	Northwest Pacific	3
Amursky Bay (AmB)	43.20°N 131.92°E	2	2	2	2	2	Northwest Pacific	1.5
Ussuriysky Bay (UB)	43.07°N 131.96°E	10	7	2	10	7	Northwest Pacific	1-2
Vostok Bay (VB)	42.89°N 132.73°E	9	8	3	9	8	Northwest Pacific	4-5
Sokcho Bay (SB)	38.21°N 128.59°E	3	2	3	3	2	Northwest Pacific	intertidal
The Kuril Islands (KI)	44.70°N 147.14°E	2	2	2	2	2	Northwest Pacific	intertidal
Friday Harbor (FH)	48.52°N 123.01°W	15	12	2	18	9	Northeast Pacific	intertidal
Whiffen Spit (WS)	48.36°N 123.72°W	7	6	1	8	5	Northeast Pacific	intertidal
California (CA)	38°N 123.42°W	5	4	1	5	4	Northeast Pacific	88 - 122
Total	-	63	53	18	67	49	-	-

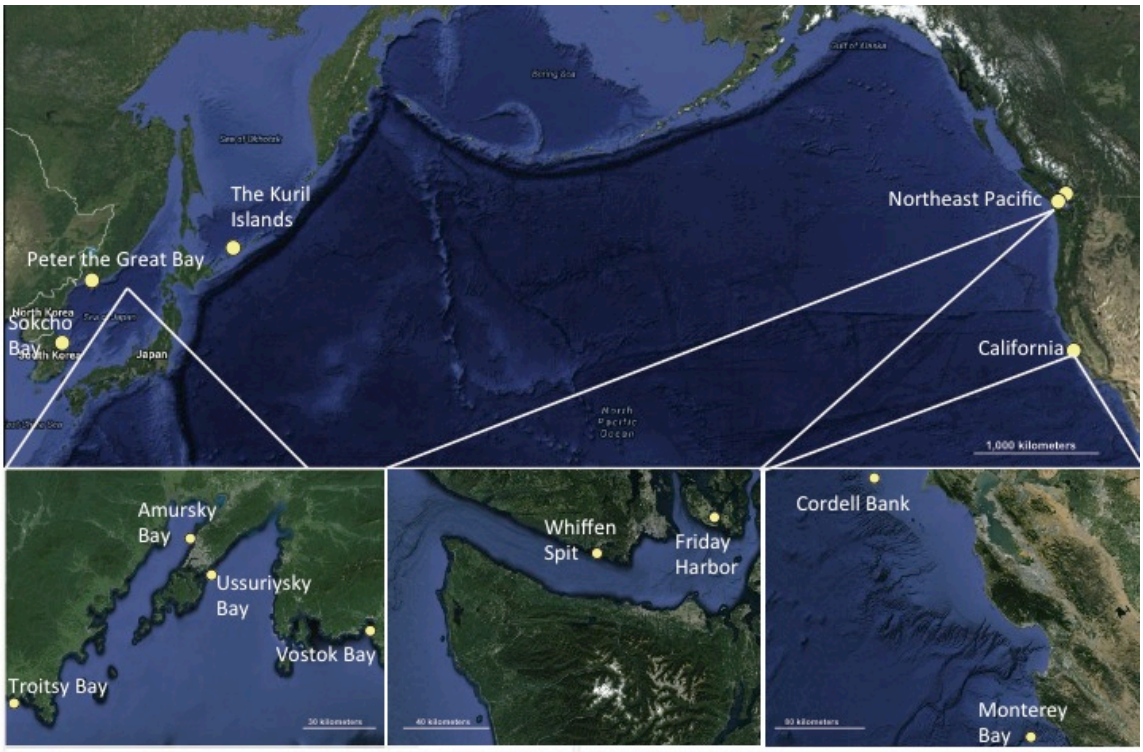

Figure 4. Map of the western Pacific and the eastern Pacific collection localities (above). Peter the Great Bay (bottom left) encompasses Vostok Bay, Amursky Bay, Ussuriysky Bay, and Troitsy Bay. The eastern Pacific (bottom center and bottom right) sampling sites were Whiffen Spit, Friday Harbor, Cordell Bank and Monterey Bay. Map data from Google, DigitalGlobe. Map edited with Adobe Illustrator.

Measures of Diversity and Genetic Distances

Mean within-population distances range from 0.2% (The Kuril Islands, Amursky Bay) to 2.6% (California) for COI for both the Tamura-Nei model and Kimura's 2-Parameter model and from 0 (Amursky Bay, Sokcho Bay) to 1.3% (California, Ussuriysky Bay) using Kimura's 2-Parameter model for 16S. Using the Tamura-Nei model for 16S, mean within-population distances ranged from 0 (Amursky Bay, Sokcho Bay) to 1.4% (Ussuriysky Bay). All standard error values were less than or equal to 4% for both datasets. Genetic distances calculated from our ITS dataset showed extremely high levels of between-coast divergence (> 1.0). This may be due to our small sample size (n = 18) or an erroneous alignment, and therefore we recommend revisiting these estimates with either a larger dataset or using a different nuclear gene (such as H3, H4, or 28S) to obtain valid and accurate estimates of genetic distance.

Mean inter-population diversity estimated using the Tamura-Nei model (1993) was 8.4% (16S) and 15.7 (COI) with a standard error of 1.3% and 1.8%, respectively. We report estimates from both models because although Tamura-Nei was selected as the most appropriate model for this data (based on results from jModelTest), Kimura's 2-Parameter model allows for an adequate comparison with other studies that have used this model for estimating population diversity. Using Kimura's 2-Parameter model, mean inter-population diversity was estimated at 8.3% (16S) and 15.3% (COI) with a standard error of 1.2% and 1.7%, respectively. Mean intra-population diversity was 0.5% (16S) and 0.9% (COI) with a standard error of 0.1% for each dataset regardless of

model selection. The coefficient of genetic differentiation (G_{ST}) when comparing eastern populations with western populations was calculated to be 0.943 (COI) and 0.945 (16S) (the table as seen on page 22). Standard error values were obtained by 1,000 bootstrap replicates.

Mean between-population genetic distances range from 0.2% to 33.1% using the Tamura-Nei model (1993) and 0.2% to 32.3% using Kimura's 2-Parameter model (1980) for the COI dataset (Table 2). For the 16S dataset, mean between-population genetic distances range from 0.1% to 17.8% using the Tamura-Nei model and 0.1% to 17.6% using Kimura's 2-Parameter model (Table 3). Mean between-coast genetic distances were 17% (16S) and 32% (COI), with standard errors of 2.5% and approximately 3.5%, respectively. These distances between populations were considerably greater than distances between same-coast populations, which were less than or equal to 2% regardless of evolutionary model or mitochondrial gene sequence.

An analysis of molecular variance gave strong support for the separation of the eastern and western populations (Table 4). Within the COI dataset, 95% of molecular variance occurred between eastern and western coasts, with -0.2% variation among populations within the groups and 4.7% variation within populations. 91.5% of total variance within the 16S dataset was attributed to variance between eastern and western coasts, with 1.15% variation within groups and 7.32% variation within populations. The greatest proportion of molecular variation that could be attributed to between-group variation was found in our COI-16S concatenated dataset (95.4%). F_{ST} values of 0.95, 0.93, and 0.96 for COI and 16S, and COI-16S respectively, provide evidence of no gene

flow occurring between eastern and western populations. The number of haplotypes recovered from these three datasets varied between 27 (16S) and 52 (COI), with 46 haplotypes being recovered in the combined dataset.

Table 2. Mean between-group genetic distances calculated for COI using the Tamura-Nei model (above diagonal) and Kimura's 2-parameter model (below diagonal). Column headings contain site name abbreviations. Values represent the average number of base substitutions per site between all groups. S.E. < 0.04 for all samples.

	TB	*AmB*	*UB*	*VB*	*SB*	*KI*	*FH*	*WS*	*CA*
Troitsy Bay		0.005	0.006	0.006	0.008	0.005	0.323	0.320	0.328
Amursky Bay	0.005		0.005	0.004	0.006	0.002	0.324	0.322	0.330
Ussuriysky Bay	0.006	0.005		0.006	0.007	0.006	0.323	0.321	0.328
Vostok Bay	0.006	0.004	0.006		0.007	0.004	0.325	0.323	0.331
Sokcho Bay	0.008	0.006	0.007	0.007		0.007	0.323	0.321	0.329
The Kuril Islands	0.005	0.002	0.006	0.004	0.007		0.323	0.321	0.327
Friday Harbor	0.315	0.317	0.316	0.318	0.316	0.316		0.013	0.020
Whiffen Spit	0.312	0.314	0.313	0.315	0.313	0.314	0.013		0.018
California	0.320	0.322	0.320	0.323	0.322	0.320	0.020	0.018	

Table 3. Mean between-group genetic distances calculated for 16S using the Tamura-Nei model (above diagonal) and Kimura's 2-parameter model (below diagonal). Column headings contain site name abbreviations. Values represent the average number of base substitutions per site between all groups. S.E. < 0.03 for all samples.

	TB	*AmB*	*UB*	*VB*	*SB*	*KI*	*FH*	*WS*	*CA*
Troitsy Bay		0.003	0.009	0.004	0.011	0.013	0.173	0.171	0.174
Amursky Bay	0.003		0.007	0.001	0.013	0.016	0.177	0.175	0.178
Ussuriysky Bay	0.009	0.007		0.008	0.014	0.016	0.174	0.172	0.174
Vostok Bay	0.004	0.001	0.008		0.014	0.017	0.177	0.175	0.178
Sokcho Bay	0.011	0.013	0.014	0.014		0.003	0.157	0.154	0.157
The Kuril Islands	0.013	0.016	0.016	0.017	0.003		0.160	0.158	0.159
Friday Harbor	0.171	0.175	0.172	0.175	0.156	0.159		0.003	0.010
Whiffen Spit	0.169	0.173	0.170	0.173	0.153	0.157	0.003		0.010
California	0.172	0.176	0.172	0.176	0.156	0.158	0.010	0.010	

Table 4. The results of our AMOVA analysis on the COI, 16S, and concatenated datasets. G_{ST} values were calculated in MEGA 5.05 using Kimura's 2-Parameter model comparing eastern and western populations and incorporating a gamma distribution of rate variation. Support for G_{ST} was calculated over 1,000 bootstrap replicates. * represents significant values (p < 0.01).

Loci	*Between Groups*	*Between Populations*	*Within Populations*	F_{CT}	F_{SC}	F_{ST}	G_{ST}	*Haplotypes (n)*
COI	95.47%	-0.20%	4.74%	0.95	-0.05	0.95*	0.933	52
16S	91.47%	1.15%	7.32%	0.92*	0.13	0.93*	0.925	27
Both	95.40%	0.34%	4.26%	0.95*	0.07	0.96*	0.926	46

Phylogeny

Phylogenetic trees were constructed using a concatenated mtDNA dataset representing 72 individuals, including five outgroup specimens (the table as seen on page 16). These phylogenies, constructed according to ML and Bayesian inference (Figure 5, as seen on page 24) recovered consistent eastern Pacific and western Pacific clades with high bootstrap support/posterior probability (100%/1.0). The same western and eastern clades were recovered using the three-gene concatenated dataset, as well as in NJ trees constructed with single-gene datasets (100% bootstrap support). This indicates that the divergence of these clades is represented in both mitochondrial and nuclear sequence data and is not reliant on a single evolutionary model. However, because ITS sequences were not available for the five outgroup specimens, it was not

possible to root the tree using our nuclear data. The presence of the eastern and western clades are independent of outgroup selection, as they remained well supported with multiple different outgroup combinations. Phylogenetic trees constructed using Bayesian inference resulted in some fine-scale genetic structure within a coastal clade with strong support (> 0.80 posterior probability), though these groupings were not associated with any geographical region and were not recovered in all analyses. Trees constructed in MEGA 5.05 using Maximum Parsimony, Minimum-Evolution, and Neighbor-Joining (the latter two trees incorporating only transversions) recovered the same eastern and western clades without any reproducible structure exhibiting high bootstrap support (< 50%).

In all phylogenetic reconstructions obtained with any combination of mtDNA sequence information (COI only, 16S only, and COI-16S), the eastern clade of *P. agassizii* and the western clade of *P. agassizii* were not recovered as sister groups. The gene phylogenies indicate that their divergence precedes the divergence of the eastern clade from *P. nigrescens*, and from the common ancestor of *P. scolops* and *P. stephensoni*. The only separation event that occurred prior to the divergence of the eastern and western clades was the divergence of *P. granulatum*, which represents the root of our tree, from the common ancestor of the rest of the species. *P. stephensoni* and *P. scolops* represent the most closely-related species to the eastern clade of *P. agassizii* The western clade, on the other hand, is most closely related to *P. perlucens* (which has since been shown to be a cryptic species complex (Kawauchi & Giribet, 2010)).

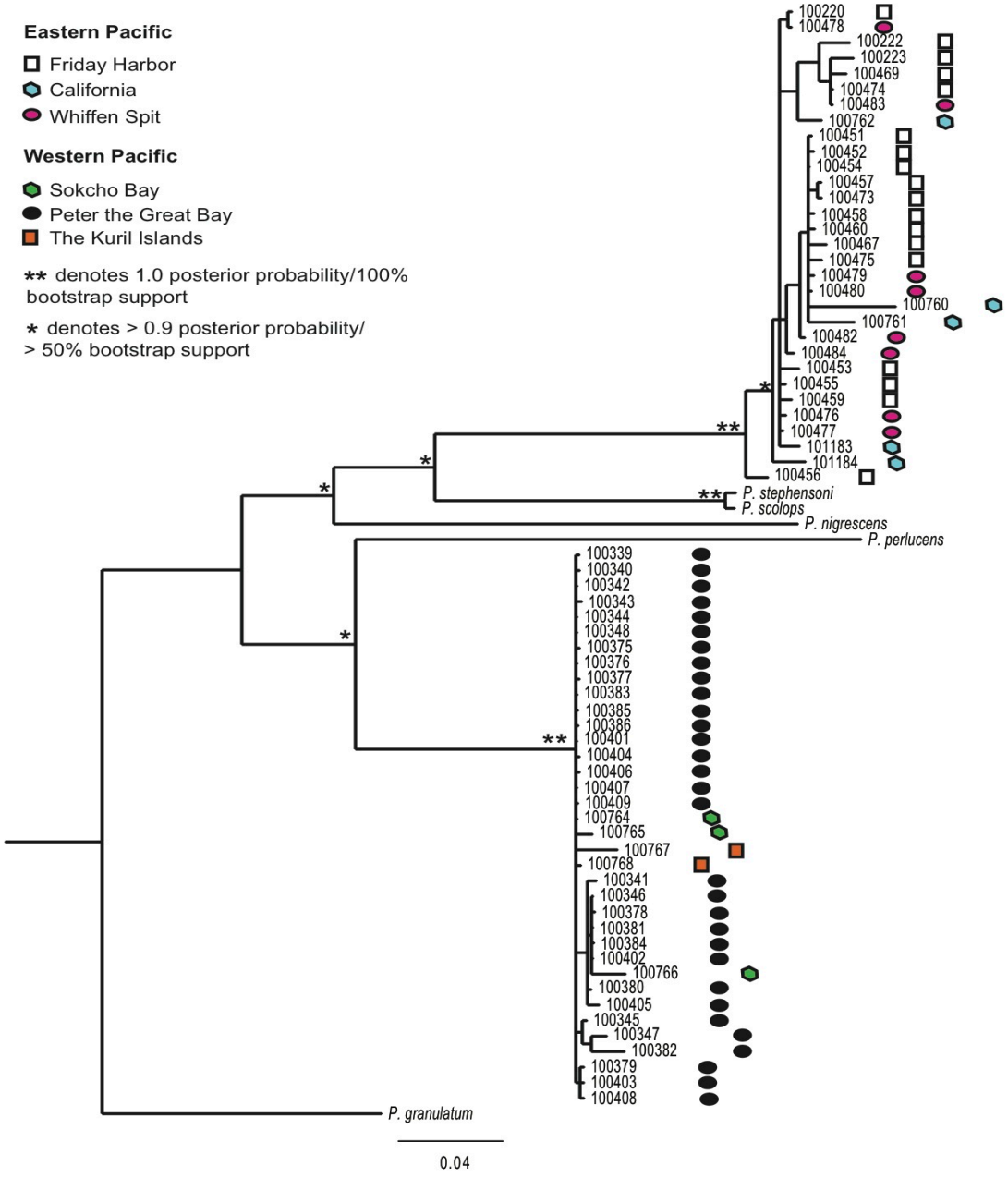

Figure 5. Phylogenetic tree based on the COI-16S sequences of 67 *P. agassizii* individuals. The eastern and western clades were recovered in all phylogenetic analyses using either NJ, ML, or BI. This tree was constructed using Bayesian Inference according to a GTR+I+G model over 10,000,000 generations with a 25% burn-in period. Bootstrap support was determined from a ML tree constructed using a GTR+I+G model with 1,000 bootstrap replicates. Outgroups were selected based on their relation to *P. agassizii* determined from previous studies.

Discussion

Our data indicates that *P. agassizii* is not monophyletic throughout its distribution range, a finding consistent with the previous study by Schulze et al. (2012). Instead, *P. agassizii* consists of two highly divergent species, with corresponding phylogeographic associations to the west and east Pacific coasts. Phylogenetic analyses (Figure 5) show that the western Pacific clade split off early in the evolution of the genus and is more closely related to *P. perlucens*. Together, these two taxa emerge as sisters to the group formed by *P. nigrescens, P. stephensoni* + *P. scolops*, and the eastern Pacific clade. Regardless of which *Phascolosoma* species is used to root our trees, the eastern and western *P. agassizii* clades are not identified as sister taxa. This early split between the lineages would explain the large genetic distances between the two *P. agassizii* clades. The general consensus for maximum intra-species distance seems to be approximately 2-3% (Herbert *et al.*, 2003). Each of the populations along a given coast separately falls within that range, but the average distance between the coasts far exceeds this value.

Our fixation indices measured for both COI and 16S, indicate that the eastern and western Pacific populations of *P. agassizii* are genetically isolated from each other. This notion is supported by our AMOVA statistics, with greater than 95% of observed molecular variance attributed to east-west divergence. The differences in coastal variation values between 16S and COI may be attributed to the faster mutation rate found in COI. This would skew estimates of genetic diversity measured for these populations, especially in the presence of fine-scale population structure. A separate

study using ISSR-PCR is currently being conducted to examine population structure within the western and eastern Pacific sample sets. This method is a rapid and inexpensive technique that categorizes fine-scale genetic structure through a size-fragment analysis of nuclear loci flanked by microsatellites. It provides genetic information on intraspecies populations or closely-related species in a shorter amount of time than DNA sequencing requires.

The presence of a long-lived pelagosphera larvae in populations on both coasts was thought to support a cosmopolitan distribution of these sipunculans (Scheltema & Hall, 1975). However, if these two populations belong to separate species, then this pelagosphera stage is most likely symplesiomorphic in these lineages. Our findings suggest that no significant genetic exchange occurs between the eastern and western species. However, this does not disprove that the larvae are capable of crossing the distance between the coasts. It is possible that they simply do not encounter suitable habitat for metamorphosis on the opposite coasts or that they occasionally do, but either do not survive to adulthood or have not been sampled in our study. Physical oceanographic modeling may provide additional insight into levels and distances of dispersal within *P. agassizii* populations, and thus the extent of connectivity between coasts and along a coastline. Pelagosphera larvae of *P. agassizii* have been kept alive in laboratory conditions for up to seven months (Rice, 1967). Although it is unknown how long this larval duration lasts in the field, a previous study of other sipunculan pelagosphera larvae suggests that the maximum lifespan in the field might be closer to three to four months. This was inferred based on current speeds and the distance from

the sampling site to the closest near-shore habitat (Scheltema & Hall, 1975). The complexity of the Pacific coastline makes it difficult to predict with much specificity areas of sipunculan recruitment. However, the genetic relationships found in this study suggest that adequate levels of dispersal and recruitment are present along the eastern and western Pacific coasts. The lack of fine-scale genetic structure associated with any particular population is intriguing, especially considering that the Sea of Japan is an almost entirely enclosed region. Our data implies that gene flow is occurring between the Kuril Island populations and those located within the Sea of Japan, despite the restricted geography of the latter basin. It is possible that westward larval dispersal is occurring via oceanic currents, transporting individuals from the Kuril Islands to suitable habitats inland. This would potentially facilitate reproduction and gene flow between these western populations.

Cutler (1994) recognized two subspecies within *P. agassizii*, *P. agassizii agassizii* (Keferstein 1866) and *P. agassizii kurilense* (Sato 1937), with the former's western Pacific distribution largely occurring in Japanese waters while the latter is found only in the Kuril Islands. Additionally, *P. agassizii kurilense* is differentiated from *P. agassizii agassizii* by the presence of a small secondary lobe on the nephridia of mature specimens (Cutler, 1994). We were unable to examine the internal anatomy of our samples from the Kuril Islands, so we cannot verify the presence or absence of this diagnostic character. Previous observations of individuals from this population have verified the presence of an expanded secondary lobe, however it should be noted that they were made using old formalin-fixed specimens in poor condition (Maiorova, pers.

obs.) Our data suggest that even if there are morphological distinctions between these two subspecies, there are no reproductive barriers between populations from the Kuril Islands and the rest of the populations within the western clade. Further, much of the current literature treats *kurilense* as a synonym of *agassizii* (Maiorova & Adrianov, 2013), and therefore the two seem to be functionally the same species.

This genetic analysis of *P. agassizii* samples collected throughout its geographic range indicates sufficient levels of divergence to taxonomically split this morphospecies complex into two isolated cryptic species. We recommend that steps be taken to revise the taxonomic information for *P. agassizii* to reflect its differentiation from the taxon found in the western Pacific, as the type specimen for the original description came from California. Further, the western Pacific populations should be formally described as corresponding to a new sipunculan species, as these do not appear to match the distributions of any species synonymized under *P. agassizii* (Cutler, 1994). Previous work has shown differences between the developmental events and environmental preferences for western Pacific individuals compared to those in the eastern Pacific, including longer spawning seasons and longer trochophore duration times, with smaller egg sizes observed in the Sea of Japan (A. Adrianov & Maiorova, 2010; Rice, 1967; Schulze *et al.*, 2012). These developmental trends were thought to be a result of phenotypic plasticity brought about by varying water temperatures, since temperature fluctuations are greater in the eastern Pacific localities than in the relatively stable western Pacific localities. Although we cannot rule out effects of phenotypic plasticity on the pelagic larval duration in sipunculans, given the high degree of genetic

divergence between the two species, the differences in larval development probably have a genetic basis as well.

A formal revision of *P. agassizii* and the description of the western Pacific species should also incorporate an in-depth morphological analysis to determine whether they represent truly cryptic species or whether morphological differences exist which are not reflected in their taxonomy (making them pseudo-cryptic or pseudo-sibling species). The shape and characteristics of proteinaceous hooks located on the anterior portion of the introvert are the most commonly used species-specific structures for identifying *Phascolosoma* species (Cutler, 1994). Schulze et al. (2012) presented microscopic images of the hooks from both species and although both fit the description of *P. agassizii*, they show notable differences in shape and proportions (the figure as seen on page 12). However, hook shape and size can vary depending on the size of the specimen and the position along the introvert. To thoroughly evaluate this feature, carefully preserved material with fully extended introverts would be necessary. The samples used for this study were largely contracted and unsuitable for morphometric analyses. Though these eastern and western clades should currently be considered cryptic species, it is not unusual for diagnostic morphological features to be found upon reanalyzing a detected cryptic species complex (Nygren, 2013).

The sparse fossil record of sipunculans makes estimations of evolutionary timescales difficult, however recent analyses using relaxed molecular clocks help to reduce uncertainty in phylogenetic dating (Drummond *et al.*, 2006; Drummond & Rambaut, 2007). Though we cannot currently predict a sufficient divergence event for

the separation of *P. agassizii*, further studies incorporating relaxed molecular clocks compared with a literature review of COI mutation rates within annelids may prove to be useful starting points to determine the evolutionary history of *Phascolosoma*.

Our current study has provided sufficient genetic data to suggest that *Phascolosoma agassizii* includes at least two cryptic allopatric lineages, with amphi-Pacific distribution, comprising the majority of this complex's geographic range. Thus, *P. agassizii* appears to be the latest member of a growing number of nominal sipunculan species that are in fact species complexes (Kawauchi & Giribet, 2010; Kawauchi & Giribet, 2013; Schulze *et al.*, 2012; Staton & Rice, 1999). Although in some of these complexes morphological differences among the lineages have been reported in hindsight (Kawauchi & Giribet, 2010), genetic identification methods provide the best initial detection of these divergent lineages. Certain estimates put the level of cryptic diversity within sipunculans at greater than 50% (Appeltans *et al.*, 2012). Whether this is an accurate representation or not, it is clear that the current number of 149 morphologically-distinguishable sipunculan species underestimates the true level of diversity within this group. Aside from their practical taxonomic benefits, studies such as this highlight the gaps in our knowledge of speciation events and provide baseline data for follow-up studies regarding larval dispersal, speciation mechanisms, and changes in species distributions. Further, with the current threats facing the health of our oceans and their biological communities, it is crucial that we accurately estimate and monitor marine biodiversity so that we can best protect these ecosystems and the services they provide.

CHAPTER III

POPULATION STRUCTURE OF *PHASCOLOSOMA AGASSIZII* DETECTED WITH ISSR-PCR AND ITS IMPLICATIONS FOR LARVAL DISPERSAL

Introduction

Phascolosoma agassizii is a well-studied marine worm in the Sipuncula clade. Though traditionally considered a phylum, Sipuncula is currently accepted as an annelid clade (Weigert *et al.*, 2014). Sipunculans represent a unique opportunity to study population connectivity because this relatively small taxonomic group (149 species) exhibits a variety of developmental modes, from direct development to possessing a long-lived planktotrophic stage called a pelagosphera larvae. This diversity of life-histories can therefore be used to further understand the relationship between larval duration, realized larval dispersal, and gene flow among adult populations.

The sipunculan *P. agassizii*, which possesses a long-lived planktotrophic pelagosphera form, was thought to have an amphi-Pacific distribution (Cutler, 1994). However, a phylogeny of Pacific samples of this species found a high degree of genetic differentiation between coastal populations (Schulze *et al.*, 2012). An additional study, with greater sample coverage in the eastern and western coasts of the North Pacific, showed no gene flow occurring between coasts as the samples from either coast belonged to reciprocally monophyletic clades (Chapter II). Therefore, *P. agassizii* comprises at least two cryptic or pseudo-cryptic species. The ontogenetic development of specimens belonging to these two clades of *P. agassizii* is well documented. In

laboratory conditions, eastern Pacific *P. agassizii* pelagospherae have been kept alive for approximately seven months (Rice, 1967); in the western Pacific, it is thought that larvae settle and metamorphose after approximately one month (A. Adrianov & Maiorova, 2010). Adults of *P. agassizii* in both eastern and western Pacific populations are benthic and semi-sessile, and thus the majority of their dispersal occurs during their larval stage. Although the presence of a long-lived planktonic larval form often predicts high dispersal and low genetic structure (in the absence of physical barriers), this is not always the case (Kawauchi & Giribet, 2013; Knowlton & Keller, 1986; Levin, 2006 and references therein).

Populations of the western clade are highly abundant within the Sea of Japan, from Peter the Great Bay off the southern coast of Russia to South Korea's eastern coast. Additionally, members of this clade can also be found in the Kuril Islands northeast of Japan. The Sea of Japan, and specifically Peter the Great Bay, is a relatively enclosed basin with a few straits along its eastern and southern borders allowing the entry of Pacific currents. Conversely, the Kuril Islands form a long archipelago on the edge of the western Pacific continental shelf, and are subject to currents from the North Pacific Gyre, specifically the Oyashio Current (Nishioka *et al.*, 2007). This stream of cool waters originates in the Arctic and stretches south to eastern Japan, where it collides with the northward Kuroshio Current and redirects eastward. These differing environmental regimes may potentially influence the larval dispersal of *Phascolosoma agassizii* along the west Pacific coast, and therefore could have a significant impact on population connectivity and gene flow. Populations of *P. agassizii* from the eastern Pacific can be

found from Alaska to Mexico (Cutler, 1994). The samples for this study were collected from Vancouver, the San Juan Islands, and California. Compared to those in the western Pacific, the oceanography of this region is more homogeneous, with the Strait of Juan de Fuca allowing for potential connectivity between populations off the U.S. coast and those from the more inland San Juan Islands.

Schulze et al. (2012) and our study from Chapter II used sequence data from two mitochondrial genes to show the presence of two cryptic species within *P. agassizii.* These clades were isolated to the western Pacific coast and eastern Pacific coast, respectively. An examination of the nuclear ITS gene of eighteen individuals also showed two coastal clades within this nominal species. However, the levels of coastal variation within the mitochondrial and nuclear sequences were low and did not provide evidence of within-region population differentiation (Chapter II). Therefore, in order to answer questions regarding realized larval dispersal and population connectivity, a more variable genetic marker is needed to properly test whether fine-scale genetic structure exists.

Inter Simple-Sequence Repeat Polymerase Chain Reaction (ISSR-PCR) is a method of fragment-length polymorphism analysis that amplifies nuclear loci flanked by microsatellites. This technique represents a rapid and inexpensive method for analyzing fine-scale population structure within a single species or between closely related species (Coupé *et al.*, 2011; Zietkiewicz *et al.*, 1994). A single primer is designed to attach to regions containing a repeated sequence followed by a unique sequence, thereby allowing for amplification in both the 5' and 3' direction. This technique amplifies numerous

variable regions within the nuclear genome simultaneously, which are then run through a gel electrophoresis protocol and visualized. Thus, it allows for the detection of more subtle genetic differentiation than DNA sequencing methods. In contrast to traditional microsatellite approaches, no prior knowledge of specific DNA sequences is required for ISSR analysis (Ye *et al.*, 2012). Amplified regions are typically non-coding, ensuring selective-neutrality of the loci being scored (Demarchi *et al.*, 2010). The relatively longer bands obtained using ISSR and the higher annealing temperatures associated with this technique translate to greater repeatability compared to many approaches using traditional RAPD primers (Semagn *et al.*, 2006). Additionally, it generates more amplified polymorphic loci per sample than RAPDs (Zietkiewicz *et al.*, 1994) and can be developed and scored faster than most AFLPs (De Aranzamendi *et al.*, 2009).

In this study, we used ISSR-PCR to amplify multiple nuclear loci within ten populations of *Phascolosoma agassizii* (Table 5). We then used these polymorphisms to compute estimates of heterozygosity within each population and gene flow between populations. Because this technique is best suited for closely-related individuals or intraspecific comparisons, we restricted our analysis to same-coast populations; this study did not compare the genetic divergence between eastern and western Pacific populations. The data gathered via ISSR-PCR was used to determine whether populations along a coastline are interbreeding or if some regions contain genetically distinct populations.

Methods

Specimen Collection

Samples were collected from a total of nine different populations on both the eastern and western Pacific coasts (Table 5). Amursky Bay, Troitsy Bay, Ussuriysky Bay, and Vostok Bay are all located off the coast of Russia in Peter the Great Bay. Individuals from Sokcho Bay, South Korea and Iturup Island in the Kuril Islands completed the western Pacific sample set. Samples from Friday Harbor (WA), Whiffen Spit (BC), Cordell Bank, and Monterey Bay (CA) populations comprised the eastern Pacific sample set. Samples from Peter the Great Bay, Friday Harbor, and Whiffen Spit, B.C., were from previous collections by Anja Schulze and colleagues. Christina Piotrowski at the California Academy of Sciences provided specimens from Cordell Bank and Monterey Bay. Anastassya Maiorova at the Zhirmunsky Institute of Marine Biology, Far East Branch of the Russian Academy of Sciences provided specimens from Sokcho Bay and the Kuril Islands. Individuals were either hand-collected by SCUBA divers or collected via bottom trawling by research vessels. Except for three specimens collected at 88 meters in Cordell Bank and two specimens collected at 122 meters in Monterey Bay, all were collected from the shallow intertidal or subtidal regions of their associated locales. All were fixed in 95% ethanol and kept at-20° C.

Table 5. Sampling localities represented in this study along with their respective sample sizes, locations within the Pacific, and collection depths (if available).

Population	*Sample Size*	*Pacific Coast*	*Depth*
Troitsy Bay	8	West	3 m
Amursky Bay	2	West	1.5 m
Ussuriysky Bay	7	West	1 – 2 m
Vostok Bay	7	West	4 – 5 m
Sokcho Bay	3	West	-
The Kuril Islands	2	West	-
Whiffen Spit	6	East	intertidal
Friday Harbor	10	East	intertidal
Cordell Bank	3	East	86 m
Monterey Bay	2	East	122 m

DNA Extraction and ISSR Amplification

DNA was extracted from the retractor muscles of individuals according to the Qiagen DNeasy Blood and Tissue Kit® protocol. If a specimen was too small (< 10 mm body length) to obtain a sample of the retractor muscle, DNA was extracted from the body wall. Quantity and quality of DNA was then measured using a Nanodrop 2000

Spectrophotometer (Thermo Scientific) to ensure similar and adequate concentrations were obtainable with this extraction method. 1 µl template DNA (0.22 pMol – 0.35 pMol original concentration) was combined with 12.5 µl Gotaq® Green Master Mix (Promega), $MgCl_2$ (2.5 mM), and a given primer (0.4x mM), in a total volume of 25 µl to create a thermocycling master mix. Amplification of ISSR regions was performed with four primer combinations (the table as seen on page 42), using the following thermocycling protocol: an initial denaturation step at 94° C for 3 min, a cycle of denaturation at 94° C for 30 s, annealing at 48° C for 30 s, and extension at 72° C for 45 s repeated 35 times, and a final extension at 72° C for 5 min. Amplified samples were visualized on a 1.0% Agarose gel. The gel was run at 80 volts (400 mA) for 2 hours and 40 minutes and stained in 10X EtBr solution for 2 hours to ensure adequate visualization of bands. Bands were visualized and scored on a Bio-Rad Molecular Imager® Gel Doc™ XR System using Quantity One© 4.6 1-D Analysis software. Each sample was amplified at least twice separately to ensure bands were reproducible (the figure as seen on page 41). Bands present in only a single amplification of a specific sample were excluded from analyses, and a third amplification was done to ensure consistent band scoring. To ensure reliable estimates of population structure, individuals that exhibited consistently smeared or absent bands for a given primer were eliminated from our combined dataset.

Population Analysis

Presence/absence matrices were created for each primer, consisting of all amplified samples with representation from each of the ten populations. A successfully

amplified locus, indicated by the presence of a band, represents a dominant state at that locus. The absence of a polymorphic band at a specific locus, therefore, represents the recessive state at that locus. The presence of a given polymorphic band was determined if a band of the same size as the reference band was visible within the sample's lane. Faint bands were scored as present, as long as they were visible in subsequent, separate amplifications. Streaked bands were eliminated from our analysis. After a presence/absence matrix was constructed, each primer sample set was analyzed using AFLP Surv 1.0 (Vekemans *et al.*, 2002). A locus within a sample was scored as a 1 if the band was present, a 0 if the band was absent, or a 9 if the locus was not scored. Analyses consisted of calculating values of expected heterozygosity, gene diversity, and F_{ST} for eastern and western populations separately. F_{ST} values were estimated using 10,000 permutations. Genetic distances were estimated over 1,000 bootstrap replicates. Allele frequencies were calculated assuming a Bayesian method with non-uniform prior distribution, and the populations analyzed were assumed to be in Hardy-Weinberg equilibrium. An initial analysis of the combined dataset was run according to the previously described parameters in AFLP Surv 1.0. This analysis, which measured eastern and western populations separately, assumed that each sampling locality represented a genetically-distinct population. Subsequent analyses using the same parameters were run to delineate actual genetically-distinct groups, either consisting of single populations or a combination of populations. Populations with F_{ST} values of less than 0.05, which was selected based on interpretations of similar studies (Casu *et al.*,

2006; Coupé *et al.*, 2011) or with non-significant values of F_{ST} (within a 99% confidence interval), were assumed to be a single gene pool.

STRUCTURE v. 2 (Pritchard *et al.*, 2000) was then used to determine the number of distinct clusters within our western and eastern Pacific dataset. Our dataset was prepared following the protocol in Vekemans et al. (2002), which treats the data as haploid. Each specific locus within each sample was scored as present (indicated by a 1) or absent (indicated by a 2). The second gene for each sample was scored as -9 for all loci, indicating missing data resulting from the dominant nature of the markers (Vekemans *et al.*, 2002). A burn-in period of 100,000 reps was determined to be adequate for this analysis, followed by a run-length of 300,000 subsequent repetitions. After this time the summary statistics, such as alpha, *F*, and the parameter values of P and Q, converged on a small range of values. This convergence can be interpreted as indicating a suitable run length (Pritchard *et al.*, 2003). Four iterations were used for each K. Because ISSR analyses assume present bands represent dominant loci, and that absent bands represent the homozygous recessive allele, some genotypic information cannot be directly visualized on a gel (i.e. distinctions between heterozygous individuals and homozygous dominant individuals). AFLP-Surv formally corrects for this in its estimates of genetic diversity, however STRUCTURE only does so under the no-admixture model (Pritchard *et al.*, 2003). Despite the fact that the admixture model does not formally correct for this loss of genotypic information, the authors of STRUCTURE believe this model will provide reasonably unbiased estimates of population composition (Pritchard *et al.*, 2003). Because of its ability to account for mixed ancestry and hybrid

zones, we decided to compare the admixture model's results with those under the no-admixture model. We ran both admixture and no-admixture models for each dataset, incorporating a correlated allele frequency model and using our sampling locations as prior information for clustering (Falush *et al.*, 2003). We used STRUCTURE Harvester (Earl, 2012) to interpret the results and determine the optimal number of populations (K) according to the Evanno method (2005). An analysis of molecular variance (AMOVA) was run in the program Arlequin v 3.5 (Excoffier & Lischer, 2010) to determine the variation between groups, among groups, and between populations. Groups were determined to be populations or combinations of populations that our analysis determined to be genetically unique. Thus, they either exhibited F_{ST} values greater than or equal to 0.05 or were detected as a unique cluster in STRUCTURE.

Results

Polymorphic Loci

A combined matrix constructed from four different primers for both the eastern and western populations resulted in a dataset of 62 scored polymorphic loci, comprised of 21 eastern individuals from four localities and 29 western individuals from six localities. Visualized bands ranged from 350 bp to roughly 1500 bp in length for each of the four primers used (Figure 6). The number of polymorphic bands per lane obtained from each primer ranged from two to thirteen, with an average of seven bands per individual sample per primer (Table 6). A significant F_{ST} estimate of 0.11 was obtained among the western populations, and a non-significant F_{ST} estimate of 0.09 was obtained among the eastern populations (assuming each locality represented a genetically-distinct

population). Estimates of total gene diversity (H_T) for the eastern and western populations were 0.345 and 0.379 respectively, with estimates of within-population diversity (H_S) equaling 0.313 and 0.338. Estimates of Nei's D_{ST}, therefore, were found to be 0.03 and 0.04 for the eastern and western populations, respectively.

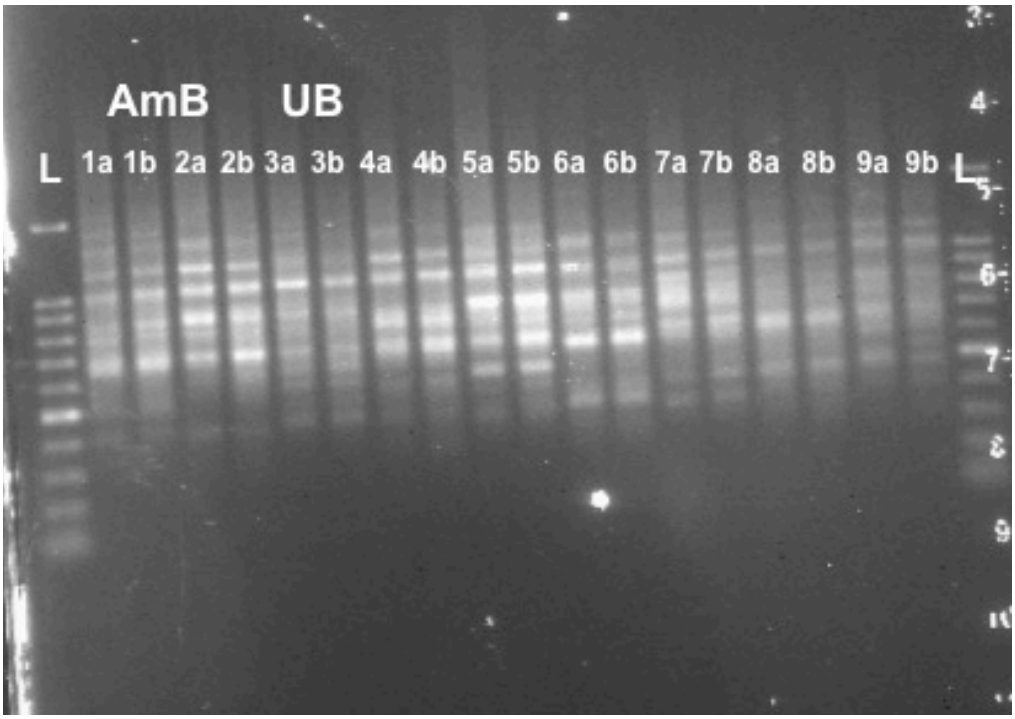

Figure 6. Image of a 1.0% TBE agarose gel stained with EtBr and run for 140 min at 80 volts. Specimens from Amursky Bay (1-2) and Ussuriysky Bay (3-9) were amplified twice (a and b) to ensure accurate band identification using primer II ($(AC)_8CA$). Ladders (L) are 5 PRIME PerfectSize 100 bp XL ladders, with 11 fragments ranging from1,500 bp to 100 bp.

Table 6. Primer sequences, sample sizes from each coast, number of polymorphic loci scored for each primer, and the average number of dominant alleles per individual for each primer.

	Sequence	*Sample size (n)*	*# of polymorphic loci*	*Mean # of bands per lane*
Primer I West	$(GTG)_4GC$	30	16	7.43
Primer II West	$(AC)_8CA$	32	15	7.75
Primer III West	$(AG)_8G$	31	16	5.16
Primer IV West	$(AC)_8T$	30	15	6.90
Primer I East	$(GTG)_4GC$	29	16	6.62
Primer II East	$(AC)_8CA$	27	15	7.61
Primer III East	$(AG)_8G$	22	16	5.36
Primer IV East	$(AC)_8T$	23	15	5.91

Genetic Structure

Amursky Bay, Ussuriysky Bay, Vostok Bay, and the Kuril Islands were found to represent genetically-distinct populations (F_{ST} = 0.12) (Table 7). Conversely, Troitsy Bay and Sokcho Bay could not be differentiated (F_{ST} = 0.038, p-value = 0.07). Instead, our data indicate that they represent a single population separate from the rest of our western Pacific localities. Assuming Hardy-Weinberg genotypic proportions, the expected heterozygosity (H) levels were 0.39, 0.28, 0.34, 0.32, and 0.34 for Amursky Bay, Ussuriysky Bay, Vostok Bay the Kuril Islands, and the combined population of

Troitsy Bay and Sokcho Bay, respectively. Pairwise F_{ST} estimates support the scenario of five distinct populations in the west (Table 8).

Within the eastern Pacific, significant estimates of F_{ST} were obtained when comparing the northeast Pacific samples with the California samples (F_{ST} = 0.07) (Table 7). The northeast Pacific population consists of Friday Harbor and Whiffen Spit, while the separate California population is comprised of individuals from Cordell Bank and Monterey Bay. No significant structure was found within the California population (p-value = 0.20), and the F_{ST} estimate obtained when testing for structure between Friday Harbor and Whiffen Spit was low enough to consider them not genetically distinct (F_{ST} = 0.038, p-value = 0.011). Therefore, our results indicate the presence of two distinct populations within the eastern Pacific, one comprised of the northeast localities and one comprised of the California localities.

Table 7. Measures of genetic diversity obtained for two different scenarios of population structure among western and eastern Pacific communities of *P. agassizii*. The combined population in the five-population scenario is comprised of Troitsy Bay, and Sokcho Bay. The northeastern Pacific population is comprised of Whiffen Spit and Friday Harbor. The California population is comprised of Monterey Bay and Cordell Bank.

	Distinct Populations	F_{ST}	H_T	H_S	*Confidence Level*	*p-Value*
West, five populations	Ussuriysky Bay, Vostok Bay, Amursky Bay, Kuril Islands, combined population	0.12	0.38	0.34	99%	0.001
East, two populations	Northeast Pacific, California	0.07	0.34	0.31	99%	0.003

Table 8. Pairwise F_{ST} estimates for the five population scenario within the western Pacific (top) and for the two population scenario within the eastern Pacific (bottom). Measurements were obtained using 1,000 bootstrap replicates.

	CP	*AmB*	*UB*	*VB*	*KI*
Combined population	0	-	-	-	
Amursky Bay	0.0987	0	-	-	
Ussuriysky Bay	0.1013	0.1276	0	-	
Vostok Bay	0.0517	0.0934	0.1453	0	
Kuril Islands	0.1179	0.1472	0.1329	0.1705	0

	C	*NP*
California	0	-
Northeast Pacific	0.0669	0

Structure Analysis and AMOVA Testing

Within the western Pacific, the uppermost hierarchical structure obtained using the Evanno method (2005) resulted in two distinct clusters. This method estimate the number of K groups from the dataset by plotting the change in the likelihood distribution over a specified range of K. The second cluster comprised individuals from Ussuriysky Bay (approximately ≥ 80% membership), the first all other individuals (approximately ≥

60% membership) (Figure 7A, as seen on page 48). However, further structure was seen when assuming three distinct clusters, with individuals from Vostok Bay representing a distinct population (Figure 7C), but this pattern was not recovered without incorporating the locality prior information. No further structure was seen within the west when we tested for up to 15 distinct clusters. The composition of clusters did not differ between the admixture and no-admixture models.

An analysis of the eastern Pacific population resulted in the detection of two distinct clusters at the K=2 scenario following the Evanno method. Cluster one comprises four individuals, with representation from Cordell Bank, Whiffen Spit, and Friday Harbor. Cluster two is comprised of all the other individuals (Figure 7B). Thus, the structure obtained assuming K = 2 is not concordant with the geographic proximity of our sampling locations nor does it coincide with our results from our AFLP-Surv analysis, which showed no significant structure within our California samples or within our northeast Pacific samples. Further structure was again seen when assuming three distinct clusters, though this was not recovered in all iterations for K = 3. Individuals from the California populations and a single individual from Friday Harbor comprised cluster one, while the majority of individuals from Whiffen Spit and Friday Harbor made up cluster two. Cluster three is represented by Cordell Bank, Whiffen Spit and Friday Harbor, the same individuals as those from cluster one when assuming K = 2 (Figure 7D).

An analysis of molecular variance (AMOVA) for the eastern population was performed incorporating two distinct groups, the first consisting of individuals from

California (Cordell Bank and Monterey Bay populations) and the second consisting of individuals from the northeast Pacific (Friday Harbor and Whiffen Spit populations). This analysis resulted in approximately 14% of the variation being attributed to variation between California samples and northeast Pacific samples (Table 9, as seen on page 49). However, the majority of variation was seen within populations (approximately 77%). Within the western Pacific, under the five-population scenario only 8.5% of the variation could be attributed to between groups. The five-population scenario was not supported by STRUCTURE, since Amursky Bay and the Kuril Islands were not identified as representing unique clusters. However, even less variation (approximately 2.5%) was seen among groups when Amursky Bay and the Kuril Islands were combined into a population with Troitsy Bay and Sokcho Bay (also supported by STRUCTURE). This program did identify Vostok Bay as a unique cluster, despite relatively low genetic differentiation with the combined western population. When Vostok Bay was grouped with this population, the among-group variation was calculated at 5.3%. When all of the locations except Ussuriysky Bay were grouped together, the among-group variation dropped to 4.7%. Therefore, we posit that the five-population scenario accounts for most genetic structure among our sampling localities within the western Pacific.

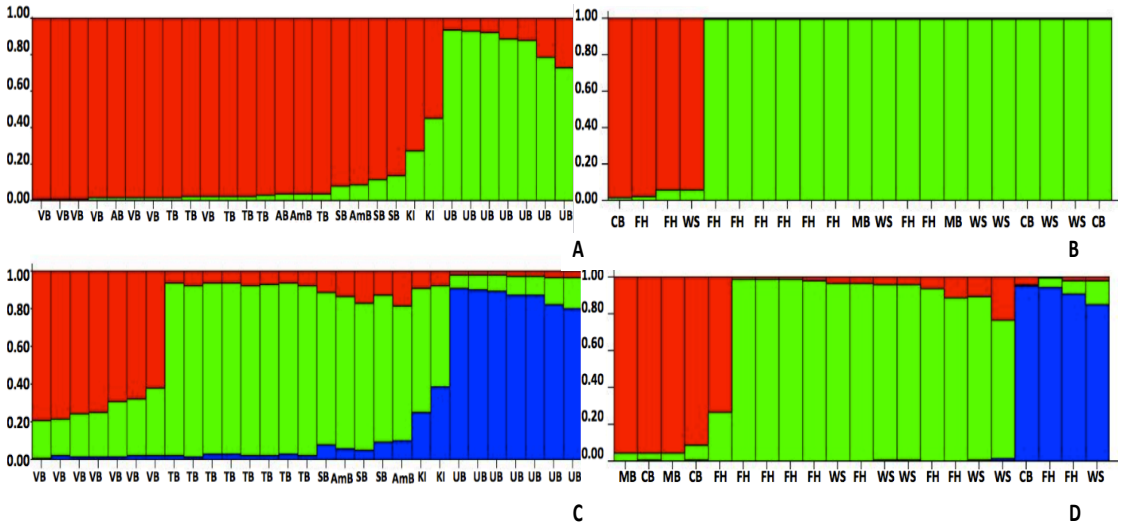

Figure 7. $\widehat{Q}$ plots representing the membership coefficients of each individual to a specific cluster. Clockwise from top left: A = Western Pacific, K = 2; B = Eastern Pacific, K = 2; C = Western Pacific, K = 3; D = Eastern Pacific, K = 3. Abbreviations on X-axis indicate sampling localities. Plots are taken from STRUCTURE runs incorporating an admixture model and using prior sampling location information. The same composition of each cluster for each K is seen using a no-admixture model.

Table 9. Three – level AMOVA testing for the eastern and western datasets. The two populations in the eastern Pacific are comprised of the California samples (Monterey Bay and Cordell Bank) and the northeast Pacific samples (Whiffen Spit and Friday Harbor). The five populations in the western Pacific are Vostok Bay, Ussuriysky Bay, Amursky Bay, the Kuril Islands, and a combined population of the Troitsy Bay and Sokcho Bay samples. F_{CT} = variance between regions relative to total variance; F_{SC} = variance among populations within regions; F_{ST} = variance among populations relative to total variance. Asterisks indicate statistically significant fixation indices, with p-value < 0.001.

Analysis of Molecular Variance	*% Variation*
East Pacific – Two Populations	-
Among groups	14.13
Among populations within groups	8.85
Within populations	77.03
F_{CT}	0.141
F_{SC}	0.103*
F_{ST}	0.230*
West Pacific – Five Populations	-
Among groups	8.54
Among populations within groups	10.78
Within populations	80.68
F_{CT}	0.085
F_{SC}	0.118
F_{ST}	0.193*

Discussion

Our data indicate the presence of five distinct populations of *Phascolosoma agassizii* within the western Pacific based on our AFLP-Surv analysis. The results from our STRUCTURE analysis, however, indicate the presence of only three distinct regions in the western Pacific. Although they were not detected as unique clusters in STRUCTURE, the relatively large pairwise F_{ST} values for Amursky Bay and the Kuril Islands suggest that these populations experience barriers to gene flow with the rest of the Sea of Japan. Additionally, the Ussuriysky Bay population exhibits significant genetic differentiation when compared with other populations in the western Pacific. Vostok Bay represents a genetically-distinct population with some degree of connectivity to our other sample sites. This genetic structure could be maintained by oceanic conditions within the region. The Sea of Japan's currents flow in a counter-clockwise direction, with the western edge dominated by the southern-moving Liman Current (Martin & Kawase, 1998). This could potentially facilitate dispersal of planktonic *P. agassizii* larvae among the coastal populations throughout the Sea of Japan, specifically between Vostok Bay, Troitsy Bay, and Sokcho Bay. However, the complex geographical structure of Amursky Bay and Ussuriysky Bay may prevent significant dispersal of larvae into these inlets. Additionally, the location of the Kuril Islands could prevent connectivity between this region and the Sea of Japan, since these islands are subject to differing currents and are connected only through the narrow La Perouse and Tsugaru Straits.

Amursky Bay and Ussuriysky Bay surround Vladivostok, a heavily populated city known for its shipping industry. Thus, the fauna of these bays experience significant anthropogenic stressors and pollution resulting from the heavy ship traffic. The genetic structure observed between Amursky Bay and Ussuriysky Bay may be a result of strong levels of local recruitment within these habitats. Modeling studies incorporating larval mortality show higher-than-expected levels of retention (Cowen *et al.*, 2000; Ellien *et al.*, 2004). In a heavily industrialized bay system, one might expect to see more larval mortality causing reduced long-distance dispersal and relatively more local recruitment. Further, large fluctuations in salinity and temperature dominate the hydrological regimes within these bays (Kuzmin *et al.*, 2001). These factors can affect larval settlement, particularly of benthic fauna (Kingsford *et al.*, 2002). Therefore, ecological differences between our sampling sites may play a role in influencing the population connectivity between the waters of Vladivostok and the rest of the Sea of Japan.

We see relatively weak structure among the eastern Pacific sampling sites. Our AFLP-Surv and STRUCTURE analysis show low, yet significant levels of genetic differentiation between the northeast populations and the California populations. Whiffen Spit and Friday Harbor are located along the Strait of Juan de Fuca, which experiences strong eastward winds in the summer season. This may prevent significant larval dispersal from these habitats westward into the Pacific Ocean. However, if larvae do exit the Strait of Juan de Fuca via the western opening, they could be transported north by the poleward-flowing Vancouver Island Coastal Current (Masson & Cummins,

1999) or south by the seasonal shelf currents (MacFadyen *et al.*, 2005). Based on our analyses and the oceanographic regimes of the northeastern Pacific, it would seem that while larval dispersal between California populations and northeastern populations can occur, it is limited. The mean annual water temperatures for Cordell Bank and Monterey Bay range from 8 - 16° C, while in the Strait of Juan de Fuca they range from 8 – 12° C (Schulze *et al.*, 2012). Additionally, the freshwater influence within the strait increases salinity fluctuations of this region relative to the offshore California waters. Therefore, it is possible that ecological differences may negatively influence larval dispersal and settlement between these populations.

In addition to oceanic currents and environmental conditions, larval behavior can be a significant factor affecting dispersal. Sipunculan trochophore larvae are planktonic and have been shown to exhibit positive phototaxis (Cutler, 1994). However, the pelagosphera larval form of *P. agassizii* is predominantly a bottom feeder, at least in laboratory conditions (Rice, 1973). During metamorphosis from the trochophore stage to the pelagosphera stage, *P. agassizii* also forms a posterior terminal organ, which is used for attaching itself to the substrate as it feeds (Rice, 1973). Thus, it appears that in addition to a morphological transformation, *P. agassizii* also undergoes an ecological shift between larval forms. While largely planktonic during its trochophore stage, it seems to exhibit benthic feeding and a preference for benthic habitats as a pelagosphera. A study of vertical migration in damselfish larvae showed that downward movement throughout ontogeny results in higher natal retention (Paris & Cowen, 2004). If *P.*

agassizii undergoes similar vertical migrations throughout its larval duration, it may display similar patterns of limited dispersal.

Adult populations of certain marine invertebrate groups are known to influence local recruitment of their larvae, and a plankton survey within Vostok Bay gave evidence for natal retention of polychaetes within this system (Omelyanenko & Kulikova, 2011). Rice (1986) showed that larvae of the sipunculan *Golfingia misakiana* (now classified as *Apionsoma misakianum*) settled and metamorphosed more frequently in areas that contained adult members of the same species. She determined that a water-soluble, low molecular weight metamorphosis-inducing factor released by adult individuals, possibly in conjunction with a specific type of substrate, was responsible for this increased settlement (Rice, 1986). Although the possibility of settlement cues for *P. agassizii* are not documented in the literature, the presence of local adult populations may positively influence local recruitment of planktonic larvae.

This study represents the first attempt to describe genetic connectivity among sipunculans at the population level using ISSR-PCR; previous studies on population structure within this group focused on mtDNA sequence data and allozyme variation (Du *et al.*, 2009; Staton & Rice, 1999; Xiaodong *et al.*, 2008). For some collecting localities, our sample sizes were extremely small and therefore may not represent the entire local population. Additional collections from the Kuril Islands, Sokcho Bay, Amursky Bay, and California would allow for a more thorough examination of intra and interpopulation diversity. The small sample sizes from Amursky Bay and the Kuril Islands may be the reason for the disparity between STRUCTURE and AFLP-Surv

regarding the populations as genetically-distinct regions. Larval transport exhibits its greatest influence on genetic structure on scales of approximately hundreds of kilometers (Hellberg *et al.*, 2002; Palumbi, 2003). However, we cannot rule out the possibility of climatic, geological, or bottleneck events affecting gene flow among these populations. Increased sample sizes and studies incorporating other genetic markers would be useful to examine this question. Records exist of *P. agassizii* in Baja California, Alaska, and the Atlantic and Indian Ocean as well (Cutler, 1994). Samples from these locations will allow for a more holistic view of this species' dispersal and population structure. However, given the reported cryptic diversity of this species within the Pacific (Chapter II, Schulze *et al.*, 2012), it is possible that these records are a result of misidentifications or represent additional cryptic lineages. Regardless, their investigation would be necessary to understand the distribution range of the nominal *P. agassizii* species.

Increasing the number of scored polymorphic loci may also contribute to a more accurate estimate of population structure. When estimating the number of distinct groups in STRUCTURE using AFLPs or ISSR data, Evanno recommends incorporating at least 100 loci to ensure the detection of all present groups (Evanno *et al.*, 2005). When comparing the results obtained using the admixture model with those obtained using the no-admixture model, we found no significant differences in the composition of recovered clusters or the estimates of membership coefficients. Thus, we believe the admixture model to be a reliable predictor of population structure in this study.

The long-held paradigm that larvae are simply passive dispersers is shifting. Studies over the past decade have shown that larval behavior plays a crucial role in their

dispersal (Krug & Zimmer, 2004; Levin & Huggett, 1990). There is a growing body of evidence that suggests longer planktonic larval durations do not necessarily mean greater connectivity among populations (Selkoe & Toonen, 2011; Shanks, 2009; Swearer *et al.*, 2002), even within Sipuncula (Staton & Rice, 1999). In this examination of *P. agassizii* population structure, we see significant restriction of gene flow between Ussuriysky Bay and the Sea of Japan, regions that are separated by less than 100 km. In addition to highlighting the gaps in our knowledge of sipunculan larval dispersal, this study brings up an issue of population management within Ussuriysky Bay. Though there are no known estimates of population size, *P. agassizii* is a common member of the benthic invertebrate community around Vladivostok and is the most common sipunculan in the Sea of Japan (A. V. Adrianov *et al.*, 2011). Therefore, it represents an important source of food for bottom-feeding fish, crabs, and other large invertebrates (A. Adrianov & Maiorova, 2010). Additionally, sipunculans play a significant role in the geochemistry of their ecosystems via their burrowing activity (Romero-Wetzel, 1987; Williams & Margolis, 1974). Ussuriysky Bay and Amursky Bay exhibit high amounts of water pollution due to the high shipping traffic and human activity in Vladivostok (Vashchenko, 2000). *P. agassizii* is extremely abundant within the waters surrounding Vladivostok, and therefore will not likely warrant management. However, further studies on the genetic connectivity of other marine fauna in this region may delineate similar restrictions to gene flow, which in turn may require separate management of these intraspecies populations.

Integrative studies combining physical modeling with the use of environmental and genetic markers can greatly advance our understanding of larval dispersal and its relation to genetic structure within marine ecosystems (Levin & Huggett, 1990). Additionally, the examination of lesser-studied groups such as sipunculans will result in a more thorough understanding of gene flow and evolution throughout metazoans.

CHAPTER IV

CONCLUSIONS

Our research confirms that the nominal species *Phascolosoma agassizii* indeed consists of two cryptic species forming reciprocally monophyletic but not sister clades within the genus *Phascolosoma*. DNA sequencing of two mitochondrial genes and one nuclear gene consistently resulted in a paraphyletic phylogeny of *P. agassizii* specimens from throughout their Pacific range. These clades were based on their geographical location, with one clade isolated to the western Pacific and one to the eastern Pacific. Our estimates of genetic distance and gene flow between these lineages indicate that their divergence and reproductive isolation occurred relatively early within the evolution of the genus. The eastern Pacific clade represents the true *P. agassizii* based on the type specimen's location, whereas the western Pacific clade comprises an undescribed species of *Phascolosoma*. Due to the paucity of sipunculan fossils and unknown mutation rates, we are unable to estimate a divergence time within this region. One could use the average COI mutation rate found in annelids such as polychaetes, which Nygren estimates to be approximately 1.9% per million years (2013). However, this method would not provide the accuracy that a specifically-calibrated molecular clock would; in the absence of such a clock, this general technique may provide good starting points for determining speciation events within *Phascolosoma*.

Our population-structure analysis confirmed that ISSR-PCR is a credible and effective method of detecting fine-scale genetic differentiation. The data indicate that

the western Pacific clade is further subdivided into between three and five genetically distinct populations, while the eastern clade contains two genetically-distinct regions, despite the presence of a long-lived planktotrophic larval form in both clades. This population structure is most likely a result of larval retention within a given spawning region. The exact mechanism used in their local recruitment is unknown and may differ between coastal clades but active habitat choice by the larvae may play a significant role. A logical next step would be a formal species description of the western Pacific clade, as well as a revision of the true range of *P. agassizii* (i.e. restricted to the eastern Pacific). This would entail a thorough morphological examination of both clades and a comparison with type material in museum collections. It would allow us to determine whether the currently recognized *P. agassizii* contains two cryptic species or two pseudo-cryptic species. The reported occurrences of this species in the Indian and Atlantic Oceans warrant investigation as well, especially in light of our population study's findings. These may be results of misidentification, further cryptic diversity, or species introductions.

Larval behavior studies, aimed at determining settlement cues and realized dispersal in both *P. agassizii* clades could provide insight into the mechanisms behind our observed genetic structure. Additionally, an ecological study of *P. agassizii* in the Sea of Japan, specifically Peter the Great Bay, would be prudent given the unique nature of the Ussuriysky Bay population. If they indeed represent a genetically-isolated cluster, further studies regarding this population may be warranted. Most investigations into Peter the Great Bay consist of marine biological surveys rather than genetic

comparisons. However, more of the latter would indicate whether Ussuriysky Bay and Amursky Bay harbor unique populations of other local marine fauna as well. Considering the high levels of pollution experienced in this region, anthropogenic bottlenecks may have resulted in distinct population structure of organisms that are particularly susceptible to environmental stress. Finally, since multiple instances of cryptic diversity have been observed within *Phascolosoma*, a revised phylogeny of the genus would be extremely useful as baseline data for future studies.

Genetic identification techniques can greatly benefit taxonomic studies, and should continue to be utilized in future research. As evidenced by our results, morphological characters alone do not always capture the true level of diversity within a species; in the case of cryptic species, they never do. It is critical, therefore, that genetic methods aimed at detecting diversity continue to be developed and improved upon. Interdisciplinary communication can certainly aid us in this goal, as some techniques are "adopted" by certain fields but not by others. For example, though ISSR-PCR was previously used mostly in botanical studies, recent research has shown that it can be equally effective in studies of animal diversity. An integrative taxonomy would, ideally, incorporate not only genetic and morphological information but also ecological and developmental data for a given species. This holistic classification will better equip the scientific community to detect and mitigate potential harmful effects of climate change and ocean acidification on marine environments.

REFERENCES

Adrianov, A., & Maiorova, A. (2010). Reproduction and development of common species of peanut worms (Sipuncula) from the Sea of Japan. *Russian Journal of Marine Biology*, 36, 1-15.

Adrianov, A. V., Maiorova, A. S., & Malakhov, V. V. (2011). Embryonic and larval development of the peanut worm *Phascolosoma agassizii* (Keferstein 1867) from the Sea of Japan (Sipuncula: Phascolosomatidea). *Invertebrate Reproduction & Development*, 55, 22-29.

Appeltans, W., Ahyong, S. T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., et al. (2012). The magnitude of global marine species diversity. *Current Biology*, 22, 2189-2202.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., et al. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, 22, 148-155.

Bleidorn, C. (2007). The role of character loss in phylogenetic reconstruction as exemplified for the Annelida. *Journal of Zoological Systematics and Evolutionary Research*, 45, 299-307.

Boore, J. L., & Staton, J. L. (2002). The mitochondrial genome of the sipunculid *Phascolopsis gouldii* supports its association with Annelida rather than Mollusca. *Molecular Biology and Evolution*, 19, 127-137.

Borda, E., Kudenov, J. D., Chevaldonné, P., Blake, J. A., Desbruyères, D., Fabri, M. C., et al. (2013). Cryptic species of *Archinome* (Annelida: Amphinomida) from vents and seeps. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20131876.

Casu, M., Casu, D., Lai, T., Cossu, P., & Curini-Galletti, M. (2006). Inter-simple sequence repeat markers reveal strong genetic differentiation among populations of the endangered mollusc *Patella ferruginea* (Gastropoda: Patellidae) from two Sardinian marine protected areas. *Marine Biology*, 149, 1163-1174.

Coupé, S., Couvray, S., Tarnowska, K., Simide, R., Hachfi, L., Pierre, S., et al. (2011). In situ spine sampling and Inter-Simple Sequence Repeat (ISSR) PCRs proved relevant methods for molecular studies on the sea urchin *Paracentrotus lividus*. *Journal of Experimental Marine Biology and Ecology*, 409, 310-314.

Cowen, R. K., Lwiza, K. M., Sponaugle, S., Paris, C. B., & Olson, D. B. (2000). Connectivity of marine populations: open or closed? *Science*, 287, 857-859.

Cutler, E. B. (1994). *The Sipuncula: their systematics, biology, and evolution*: New York: Cornell University Press.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85, 407-415.

De Aranzamendi, M. C., Gardenal, C. N., Martin, J. P., & Bastida, R. (2009). Limpets of the genus *Nacella* (Patellogastropoda) from the Southwestern Atlantic: species identification based on molecular data. *Journal of Molluscan Studies*, 75, 241-251.

Demarchi, M., Chiappero, M. B., Tatián, M., & Sahade, R. (2010). Population genetic structure of the Antarctic ascidian *Aplidium falklandicum* from Scotia Arc and South Shetland Islands. *Polar Biology*, 33, 1567-1576.

Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS biology*, 4, e88.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *Bmc Evolutionary Biology*, 7, 214.

Du, X., Chen, Z., Deng, Y., & Wang, Q. (2009). Comparative analysis of genetic diversity and population structure of *Sipunculus nudus* as revealed by mitochondrial COI sequences. *Biochemical Genetics*, 47, 884-891.

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

Ellien, C., Thiébaut, E., Dumas, F., Salomon, J.-C., & Nival, P. (2004). A modelling study of the respective role of hydrodynamic processes and larval mortality on larval dispersal and recruitment of benthic invertebrates: example of *Pectinaria koreni* (Annelida: Polychaeta) in the Bay of Seine (English Channel). *Journal of Plankton Research*, 26, 117-132.

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., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567.

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164, 1567-1587.

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294-299.

Hellberg, M. E., Burton, R. S., Neigel, J. E., & Palumbi, S. R. (2002). Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science*, 70, 273-290.

Herbert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, S96-99.

Hoare, K., Goldson, A., Giannasi, N., & Hughes, R. (2001). Molecular Phylogeography of the Cosmopolitan Bryozoan *Celleporella hyalina*: Cryptic Speciation? *Molecular Phylogenetics and Evolution*, 18, 488-492.

Kawauchi, G. Y., & Giribet, G. (2010). Are there true cosmopolitan sipunculan worms? A genetic variation study within *Phascolosoma perlucens* (Sipuncula, Phascolosomatidae). *Marine Biology*, 157, 1417-1431.

Kawauchi, G. Y., & Giribet, G. (2013). *Sipunculus nudus* Linnaeus, 1766 (Sipuncula): cosmopolitan or a group of pseudo−cryptic species? An integrated molecular and morphological approach. *Marine Ecology*.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.

Kingsford, M. J., Leis, J. M., Shanks, A., Lindeman, K. C., Morgan, S. G., & Pineda, J. (2002). Sensory environments, larval abilities and local self-recruitment. *Bulletin of Marine Science*, 70, 309-340.

Knowlton, N. (1986). Cryptic and sibling species among the decapod Crustacea. *Journal of Crustacean Biology*, 356-363.

Knowlton, N. (1993). Sibling species in the sea. *Annual Review of Ecology and Systematics*, 24, 189-216.

Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia*, 420, 73-90.

Knowlton, N., & Keller, B. D. (1986). Larvae which fall far short of their potential: highly localized recruitment in an alpheid shrimp with extended larval development. *Bulletin of Marine Science*, 39, 213-223.

Krug, P. J., & Zimmer, R. K. (2004). Developmental dimorphism: consequences for larval behavior and dispersal potential in a marine gastropod. *The Biological Bulletin*, 207, 233-246.

Kuzmin, Y. V., Burr, G., & Jull, A. T. (2001). Radiocarbon reservoir correction ages in the Peter the Great Gulf, Sea of Japan, and eastern coast of the Kunashir, southern Kuriles (northwestern Pacific). *Radiocarbon*, 43, 477-482.

Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948.

Lazoski, C., Sole-Cava, A. M., Boury-Esnault, N., Klautau, M., & Russo, C. A. M. (2001). Cryptic speciation in a high gene flow scenario in the oviparous marine sponge *Chondrosia reniformis*. *Marine Biology*, 139, 421-429.

Levin, L. A. (2006). Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*, 46, 282-297.

Levin, L. A., & Huggett, D. V. (1990). Implications of alternative reproductive modes for seasonality and demography in an estuarine polychaete. *Ecology*, 71, 2191-2208.

Lewis, C., & Karageorgopoulos, P. (2008). A new species of *Marphysa* (Eunicidae) from the western Cape of South Africa. *Journal of the Marine Biological Association of the UK*, 88, 277-287.

MacFadyen, A., Hickey, B. M., & Foreman, M. G. (2005). Transport of surface waters from the Juan de Fuca eddy region to the Washington coast. *Continental Shelf Research*, 25, 2008-2021.

Maiorova, A. S., & Adrianov, A. V. (2013). Peanut worms of the phylum Sipuncula from the Sea of Japan with a key to species. *Deep Sea Research Part II: Topical Studies in Oceanography*, 86, 140-147.

Maltagliati, F., Camilli, L., Lardicci, C., & Castelli, A. (2001). Evidence for morphological and genetic divergence in *Perinereis cultrifera* (Polychaeta: Nereididae) from two habitat types at Elba Island. *Journal of the Marine Biological Association of the UK*, 81, 411-414.

Martin, S., & Kawase, M. (1998). The southern flux of sea ice in the Tatarskiy Strait, Japan Sea and the generation of the Liman Current. *Journal of Marine Research*, 56, 141-155.

Masson, D., & Cummins, P. F. (1999). Numerical simulations of a buoyancy-driven coastal countercurrent off Vancouver Island. *Journal of Physical Oceanography*, 29, 418-435.

Mayr, E. (1948). The bearing of the new systematics on genetical problems; the nature of species. *Advances in Genetics*, 3b, 205-237.

Nishioka, J., Ono, T., Saito, H., Nakatsuka, T., Takeda, S., Yoshimura, T., et al. (2007). Iron supply to the western subarctic Pacific: Importance of iron export from the Sea of Okhotsk. *Journal of Geophysical Research: Oceans (1978–2012)*, 112.

Nygren, A. (2013). Cryptic polychaete diversity: a review. *Zoologica Scripta*.

Nygren, A., Eklöf, J., & Pleijel, F. (2009). Arctic-boreal sibling species of *Paranaitis* (Polychaeta, Phyllodocidae). *Marine Biology Research*, 5, 315-327.

Omelyanenko, V., & Kulikova, V. (2011). Pelagic larvae of benthic invertebrates of the Vostok Bay, Peter the Great Bay, Sea of Japan: composition, phenology, and population dynamics. *Russian Journal of Marine Biology*, 37, 7-22.

Padial, J. M., Miralles, A., De la Riva, I., & Vences, M. (2010). Review: The integrative future of taxonomy. *Frontiers in Zoology*, 7, 1-14.

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, 13, 146-158.

Paris, C. B., & Cowen, R. K. (2004). Direct evidence of a biophysical retention mechanism for coral reef fish larvae. *Limnology and Oceanography*, 49, 1964-1979.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253-1256.

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

Pritchard, J. K., Wen, W., & Falush, D. (2003). Documentation for STRUCTURE software: version 2.

Provan, J., & Maggs, C. A. (2012). Unique genetic variation at a species' rear edge is under threat from global climate change. *Proceedings of the Royal Society B: Biological Sciences*, 279, 39-47.

Rice, M. E. (1967). A comparative study of the development of *P. agassizii*, *Golfingia pugettensis*, and *Themiste pyroides* with a discussion of developmental patterns in the Sipuncula. *Ophelia*, 4, 143-171.

Rice, M. E. (1973). Morphology, behavior, and histogenesis of the pelagosphera larva of *Phascolosoma agassizii* (Sipuncula). *Smithsonian Contributions to Zoology*, 132, 1-51.

Rice, M. E. (1976). Larval development and metamorphosis in Sipuncula. *American Zoology*, 16, 563-571.

Rice, M. E. (1986). Factors influencing larval metamorphosis in *Golfingia misakiana* (Sipuncula). *Bulletin of Marine Science*, 39, 362-375.

Rice, M.E. (1993). Chapter 7: Sipuncula. In: Harrison FW and Rice ME, editors. *Microscopic Anatomy of Invertebrates, Vol. 12: Onychophora, Chilopoda, and Lesser Protostomata* (pp. 237-325).

Rice, M. E., Reichardt, H. F., Piraino, J., & Young, C. M. (2012). Reproduction, development, growth, and the length of larval life of *Phascolosoma turnerae*, a wood-dwelling deep-sea sipunculan. *Invertebrate Biology*, 131, 204-215.

Rice, M. E., & Schulze, A. (2004). Molecular and morphological species identification of pelagic sipunculan larvae. *Integrative and Comparative Biology*, 44, 740-740.

Romero-Wetzel, M. (1987). Sipunculans as inhabitants of very deep, narrow burrows in deep-sea sediments. *Marine Biology*, 96, 87-91.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539-542.

Sáez, A. G., & Lozano, E. (2005). Body doubles. *Nature*, 433, 111-111.

Scheltema, R. S., & Hall, J. R. (1975). The dispersal of pelagosphera larvae by ocean currents and the geographical distribution of sipunculans. *Proceedings of the International Symposium on the biology of the Sipuncula and Echiura*. Kotor: Naucno Delo., 103-116.

Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., & Crozier, R. H. (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Annual Review of Entomology*, 55, 421-438.

Schulze, A., Cutler, E. B., & Giribet, G. (2005). Reconstructing the phylogeny of the Sipuncula. *Hydrobiologia*, 535, 277-296.

Schulze, A., Cutler, E. B., & Giribet, G. (2007). Phylogeny of sipunculan worms: A combined analysis of four gene regions and morphology. *Molecular Phylogenetics and Evolution*, 42, 171-192.

Schulze, A., & Rice, M. E. (2009). Musculature in sipunculan worms: ontogeny and ancestral states. *Evolution & Development*, 11, 97-108.

Schulze, A., Maiorova, A., Timm, L. E., & Rice, M. E. (2012). Sipunculan Larvae and "Cosmopolitan" Species. *Integrative and Comparative Biology*, 52, 497-510.

Selkoe, K., & Toonen, R. J. (2011). Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, 436, 291-305.

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology*, 5.

Shanks, A. L. (2009). Pelagic larval duration and dispersal distance revisited. *The Biological Bulletin*, 216, 373-385.

Shields, M. A., & Kedra, M. (2009). A deep burrowing sipunculan of ecological and geochemical importance. *Deep Sea Research Part I: Oceanographic Research Papers*, 56, 2057-2064.

Solé-Cava, A., Klautau, M., Boury-Esnault, N., Borojecic, R., & Thorpe, J. (1991). Genetic evidence for cryptic speciation in allopatric populations of two cosmopolitan species of the calcareous sponge genus *Clathrina*. *Marine Biology*, 111, 381-386.

Stamatakis, A. (2014). RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*, btu033v031-btu033.

Staton, J. L., & Rice, M. E. (1999). Genetic differentiation despite teleplanic larval dispersal: Allozyme variation in sipunculans of the *Apionsoma misakianum* species-complex. *Bulletin of Marine Science*, 65, 467-480.

Stephen, A. C., & Edmonds, S. J. (1972). Phyla Sipuncula and Echiura.

Swearer, S. E., Shima, J. S., Hellberg, M. E., Thorrold, S. R., Jones, G. P., Robertson, D. R., et al. (2002). Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science*, 70, 251-271.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 10.

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.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., et al. (2004). Extinction risk from climate change. *Nature*, 427, 145-148.

Thompson, B. E. (1980). A new bathyal Sipunculan from southern California, with ecological notes. *Deep Sea Research Part A. Oceanographic Research Papers*, 27, 951-957.

Vashchenko, M. (2000). Pollution in Peter the Great Bay, Sea of Japan, and its biological consequences. *Russian Journal of Marine Biology*, 26, 155-166.

Vekemans, X., Beauwens, T., Lemaire, M., & Roldán−Ruiz, I. (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11, 139-151.

Weigert, A., Helm, C., Meyer, M., Nickel, B., Arendt, D., Hausdorf, B., et al. (2014). Illuminating the base of the annelid tree using transcriptomics. *Molecular Biology and Evolution*, msu080.

Will, K. W., Mishler, B. D., & Wheeler, Q. D. (2005). The perils of DNA barcoding and the need for integrative taxonomy. *Systematic Biology*, 54, 844-851.

Williams, J., & Margolis, S. (1974). Sipunculid burrows in coral reefs: evidence for chemical and mechanical excavation.

Xiaodong, D., Zian, C., Yuewen, D., Qingheng, W., & Ronglian, H. (2008). Genetic diversity and population structure of the peanut worm (*Sipunculus nudus*) in Southern China as inferred from mitochondrial 16S rRNA sequences. *The Israeli Journal of Aquaculture*, 60, 237.

Ye, Y.-y., Li, J.-j., Wu, C.-w., Xu, M.-y., & Guo, B.-y. (2012). Genetic analysis of mussel (*Mytilus coruscus*) populations on the coast of East China Sea revealed by ISSR-PCR markers. *Biochemical Systematics and Ecology*, 45, 1-6.

Zietkiewicz, E., Rafalski, A., & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, 20, 176-183.

APPENDIX A

Table A1. Phylogenetic ID numbers, their associated collecting localities, and GenBank accession numbers for specimens used in this thesis.

Specimen ID	*Collecting locality*	*COI*	*16S*	*ITS*
100220	Friday Harbor	KM226395	KM226454	-
100222	Friday Harbor	KM226396	KM226455	-
100223	Friday Harbor	KM226401	-	-
100339	Troitsy Bay	KM226368	KM226414	-
100340	Troitsy Bay	KM226369	KM226415	-
100341	Troitsy Bay	KM226370	KM226416	-
100342	Troitsy Bay	KM226371	KM226417	-
100343	Troitsy Bay	KM226349	KM226418	KM226465
100344	Troitsy Bay	KM226350	KM226419	KM226466
100345	Troitsy Bay	KM226372	KM226420	-
100346	Troitsy Bay	KM226373	KM226421	-
100347	Troitsy Bay	KM226374	KM226422	-
100348	Troitsy Bay	KM226375	KM226423	-
100375	Amursky Bay	KM226351	KM226424	KM226467
100376	Amursky Bay	KM226352	KM226425	KM226468
100377	Ussuriysky Bay	KM226376	KM226426	-
100378	Ussuriysky Bay	KM226377	KM226427	-
100379	Ussuriysky Bay	KM226360	KM226428	KM226476
100380	Ussuriysky Bay	KM226353	KM226429	KM226469
100381	Ussuriysky Bay	KM226378	KM226430	-
100382	Ussuriysky Bay	KM226379	KM226431	-
100383	Ussuriysky Bay	KM226380	KM226432	-
100384	Ussuriysky Bay	KM226402	-	-
100385	Ussuriysky Bay	KM226403	-	-
100386	Ussuriysky Bay	KM226404	-	-

Table A1 (Continued).

Specimen ID	*Collecting locality*	*COI*	*16S*	*ITS*
100401	Vostok Bay	KM226381	KM226433	-
100402	Vostok Bay	KM226361	KM226434	KM226477
100403	Vostok Bay	KM226354	KM226435	KM226470
100404	Vostok Bay	KM226362	KM226436	KM226478
100405	Vostok Bay	KM226382	KM226437	-
100406	Vostok Bay	KM226383	KM226438	-
100407	Vostok Bay	KM226384	KM226439	-
100408	Vostok Bay	KM226385	KM226440	-
100409	Vostok Bay	KM226405	-	-
100451	Friday Harbor	KM226386	KM226441	-
100452	Friday Harbor	KM226387	KM226442	-
100453	Friday Harbor	KM226388	KM226443	-
100454	Friday Harbor	KM226389	KM226444	-
100455	Friday Harbor	KM226356	-	KM226472
100456	Friday Harbor	KM226367	KM226412	-
100457	Friday Harbor	-	KM226413	-
100458	Friday Harbor	KM226406	-	-
100459	Friday Harbor	KM226407	-	-
100460	Friday Harbor	KM226391	KM226448	-
100467	Friday Harbor	-	KM226449	-
100469	Friday Harbor	KM226408	-	-
100473	Friday Harbor	-	KM226450	-
100474	Friday Harbor	KM226392	KM226451	-
100475	Friday Harbor	KM226357	-	KM226473
100476	Whiffen Spit	KM226390	KM226445	-
100477	Whiffen Spit	KM226358	KM226446	KM226474
100478	Whiffen Spit	-	KM226447	-
100479	Whiffen Spit	KM226411	-	-

Table A1 (Continued).

Specimen ID	*Collecting locality*	*COI*	*16S*	*ITS*
100480	Whiffen Spit	KM226393	KM226452	-
100482	Whiffen Spit	KM226397	KM226456	-
100483	Whiffen Spit	KM226409	-	-
100484	Whiffen Spit	KM226398	KM226457	-
100760	California	KM226359	KM226461	KM226475
100761	California	KM226410	-	-
100762	California	KM226394	KM226453	-
100764	Sokcho Bay	KM226364	KM226460	KM226480
100765	Sokcho Bay	KM226363	KM226462	KM226479
100766	Sokcho Bay	KM226355	-	KM226471
100767	The Kuril Islands	KM226366	KM226463	KM226482
100768	The Kuril Islands	KM226365	KM226464	KM226481
101183	California	KM226399	KM226458	-
101184	California	KM226400	KM226459	-
P. perlucens	South Africa	GU190274	GU190328	-
P. nigrescens	South Africa	DQ300141	GU230182	-
P. stephensoni	South Africa	GU230174	GU230185	-
P. scolops	South Africa	GU230173	GU230184	-
P. granulatum	Spain	DQ300138	GU230181	-