AN ABSTRACT OF THE THESIS OF

Sarah Hardisty for the degree of Master of Science in Integrative Biology presented on November 29, 2022

Title: Morphology, Sediment Microbial Analysis, and Species Distribution Modeling of a New Offshore Population of Ghost Shrimp (*Neotrypaea* sp.)

Abstract approved: _____

Sarah K. Henkel

In 2019, a multi-cohort population of the typically estuarine burrowing shrimp *Neotrypaea* sp., was discovered approximately 7 miles offshore of Newport, Oregon. Morphological analysis of the offshore population, microbial sediment sequencing, and species distribution modeling was conducted with this new offshore population. In order to identify *Neotrypaea* sp. to the species level, morphological comparisons were conducted among the offshore population and two estuarine species found in Yaquina Bay – *N. californiensis* and *N. gigas*. Previously established characteristics and novel characteristics were identified and measured on ImageJ. The majority of characteristics between *N. gigas* and the offshore *Neotrypaea* were not significantly distinctive except for the parts of the eyestalk. Furthermore, to investigate potential cues for the *Neotrypaea* to settle in this new deeper-water environment, the

sediment microbial communities were examined. Surface sediment was collected from offshore and estuary locations with and without *Neotrypaea* sp. DNA was extracted from the sediment samples and sequenced using Illumina MiSeq. Bacteria of the genus *Shewanella*, which have been shown to induce oyster and mussel settlement and metamorphosis, were found in both the estuary and offshore locations and in five times greater abundance with *Neotrypaea* than without. Species distribution models were run using both presence-absence data (Boosted Regression Trees, BRT) and presence-only data (MaxEnt). Environmental data for those models included course-grained layers ($9km^2$), intermediate layers ($2\ km^2$) and fine-grained layers ($25m^2$), as well as *in situ* data. The BRT with fine-grained layers and *in situ* data performed the best. Models predicted some areas of high probability of occurrence offshore Oregon and southern Washington, and ground-truthing was conducted at 11 locations. Over 120 *Neotrypaea* sp. were collected offshore Neskowin and Nehalem, OR.

Morphology, Sediment Microbial Analysis, and Species Distribution Modeling of a New Offshore Population of Ghost Shrimp (*Neotrypaea* sp.)

by

Sarah Hardisty

A THESIS

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Master of Science

Presented November 29, 2022

Commencement June 2023

Master of Science thesis of Sarah Hardisty presented on November 29, 2022

APPROVED:

Major Professor, representing Integrative Biology

Head of the Department of Integrative Biology

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Sarah Hardisty

ACKNOWLEDGEMENTS

This work would not be possible without my advisor, Dr. Sarah Henkel; who continuously supported my research projects through a global pandemic, furthered my scientific understanding, and fostered a collaborative lab environment. Her discovery of the offshore population prompted many questions, a few of which I am grateful to have worked on. I would also like to acknowledge other mentors including Dr. John Chapman and Dr. Brett Dumbauld, whose financial and scientific support allowed me to pursue the second chapter of this thesis. Additionally, I would like to thank Dr. Cheng Li, who assisted in clearly teaching me new data analysis methods. Many thanks to Dr. Leigh Torres, Dr. Su Sponaugle, and Samuel Chan on my committee for providing feedback and support. Thank you to my fellow lab members as well for mucking through the mud with me to slurp shrimp. I have met many interesting scientists and people during my time at Oregon State University and could not be more thankful for the adventures and skills I acquired. Lastly, I would like to thank the Hatfield Student Association and the National Science Foundation for funding my research.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

LIST OF TABLES

INTRODUCTION

Two species of burrowing Thallanasid shrimp, *Neotrypaea gigas* (Dana, 1854) and *Neotrypaea californiensis* (Dana, 1854), co-occur in tidal flats from Baja California, Mexico, to southern Alaska, USA, and are considered an estuarine or shallow coastal species. In 2019, a new population of an unknown *Neotrypaea* species was discovered 12 km offshore Newport, OR, with burrow opening densities of up to 400 per m^{-2} at a site slated for marine renewable energy development, PacWave-South (Henkel *et al*. 2022). The exact species of the offshore *Neotrypaea* as well as why this population established offshore and continues to persist are unknown. In this thesis, I used morphological measurements to compare individuals from the offshore population with the two estuary species, sediment microbial diversity to identify settlement-cuing microbes, and species distribution modeling to determine the probability that they would be present elsewhere offshore. I hypothesized firstly that the morphological characteristics of the offshore species of *Neotrypaea* sp. would not be statistically different from *N. gigas*, because of the initial similarities in eyestalk shape and the male major claw. Further, in the sediment with *Neotrypaea* sp. both offshore and in the estuary, I hypothesized there would be a greater abundance of bacteria associated with invertebrate larval settlement. Lastly, I hypothesized that highly suitable habitat would extend past the initial site offshore Newport, OR, where populations currently exist.

1.1 Species Description

Neotrypaea californiensis (formerly *Callianassa californiensis*) grow up to 7.6 cm length and are beige to pink in color (MacGinitie, G. E., 1934). *N. gigas*, as their name suggests, are larger and grow to a body length of 12.7 cm and are beige in color (Holthuis 1991). They are

detritivores and planktivores, consuming sediment and filter feeding through the water that moves through their burrows (MacGinitie and MacGinitie 1949, Powell 1974). *N. gigas* generally burrow in the mud (17.3 cm mean depth) and *N. californiensis* burrow in the sand (30.8 cm mean depth), although burrow size can change depending on substrate (Griffis and Chavez 1988, Jensen 2014). *N. gigas* burrows are shallow and branched, whereas *N. californiensis* burrows are deep and Y-shaped (Griffis and Chavez 1988). Both species are found in Yaquina Bay and throughout Oregon estuaries, although *N. gigas* is not as abundant as *N. californiensis*.

1.2 Ecological Role

These burrowing shrimp are considered ecosystem engineers, as their bioturbation influences oxygen cycling in the sediment and can disrupt macrofaunal communities. They both increase the penetration of oxygen into the sediment and consume oxygen of the benthic community (Thompson and Pritchard 1969, Witbaard and Duineveld 1989, Leduc and Pilditch 2017). Presence of *Neotrypaea* is associated with fewer benthic macrofauna species compared to other types of habitat dominated by bare mud/sand, another species of burrowing shrimp *Upogebia pugettensis*, the cordgrass *Spartina alterniflora*, or the seagrass *Zostera marina* (Ferraro and Cole 2007).

1.3 Life History

N. californiensis life history has been extensively studied by Dumbauld et al (1996) and Bird (1982). *N. californiensis* grow at a rate of 3.3-mm CL year, but this may be dependent on food supply, and live to be 4-5 years old. The period of egg brooding is 5-8 weeks, and *N. californiensis* females reach reproductive maturity at 2 (Oregon) – 3 (Washington) years (Bird et

al. 1982 and Dumbauld et al. 1996). Pelagic larval phases occur in five zoea phases between 6-8 weeks for *N. californiensis*, and they float up to 20 km offshore before returning to the estuary as post-larvae (Breckenridge and Bollens, 2010; Johnson and Gonor, 1982; Pimental, 1986, Tucker-McCrow, 1971). *N. californiensis* post-larvae return to the estuary around August – October (Washington) or June – November (Oregon) and settle in the sediments where they burrow and live out their adult lives (Dumabuld et al. 1996, Dumbauld and Bosley 2018).

1.4 References

Breckenridge, J. K., & Bollens, S. M. (2010). Biological thin layer formation: interactions between the larval decapod, *Neotrypaea californiensis*, haloclines and light. Journal of Plankton Research, 32(7), 1097-1102.

Bird, E. M. (1982). *Population dynamics of thalassinidean shrimps and community effects through sediment modification*. University of Maryland, College Park.

Dumbauld, B. R., Armstrong, D. A., & Feldman, K. L. (1996). Life-history characteristics of two sympatric thalassinidean shrimps, *Neotrypaea californiensis* and *Upogebia pugettensis*, with implications for oyster culture. *Journal of Crustacean Biology*, *16*(4), 689-708.

Dumbauld, B. R., & Bosley, K. M. (2018). Recruitment ecology of burrowing shrimps in US Pacific coast estuaries. *Estuaries and Coasts*, *41*(7), 1848-1867.

Ferraro, S. P., & Cole, F. A. (2007). Benthic macrofauna–habitat associations in Willapa Bay, Washington, USA. Estuarine, Coastal and Shelf Science, 71(3-4), 491-507.

Griffis, R. B., & Chavez, F. L. (1988). Effects of sediment type on burrows of *Callianassa californiensis* Dana and C. Gigas Dana. Journal of Experimental Marine Biology and Ecology, 117(3), 239-253.

Henkel, S. K., Revelas, E. C., Wodzicki, S., & Chapman, J. (2022). Discovery of a large offshore population of the northeast Pacific burrowing shrimp *Neotrypaea* sp. (Decapoda: Axiidea). Estuarine, Coastal and Shelf Science, 274, 107936.

Holthuis, L.B. 1991 FAO Species Catalogue. Vol. 13. Marine lobsters of the world. An annotated and illustrated catalogue of species of interest to fisheries known to date. FAO Fish. Synop. 125(13):292p. Rome: FAO.

Johnson, G. E., & Gonor, J. J. (1982). The tidal exchange of *Callianassa californiensis* (Crustacea, Decapoda) larvae between the ocean and the Salmon River estuary, Oregon. Estuarine, Coastal and Shelf Science, 14(5), 501-516.

Leduc, D.R., & Pilditch, C. A. (2017). Estimating the effect of burrowing shrimp on deep-sea sediment community oxygen consumption. PeerJ, 5, e3309.

Pimentel, G. E. (1986). Recruitment of larvae of the ghost shrimp, *Callianassa californiensis* (Crustacea, Decapoda), in the south slough estuary, Oregon [Masters Thesis, University of Oregon] Scholars' Bank at University of Oregon.

MacGinitie, G. E. (1934). The natural history of *Callianassa californiensis* Dana. American Midland Naturalist, 166-177.

McCrow, L. T. (1971). The ghost shrimp, *Callianassa californiensis* Dana, 1854, in Yaquina Bay, Oregon. . [Masters Thesis, Oregon State University] Scholars Archive @OSU

MacGinitie, G. E., & MacGinitie, N. (1949). Natural history of marine animals. USA: McGraw-Hill Book Company, Inc..

Witbaard, R., & Duineveld, G. C. A. (1989). Some aspects of the biology and ecology of the burrowing shrimp *Callianassa subterranea* (Montagu) (Thalassinidea) from the southern North Sea. Sarsia, 74(3), 209-219.

Powell, R. R./ (1974). The functional morphology of the fore-guts of the thalassinid crustaceans, *Callianassa californiensis* and *Upogebia pugettensis.* [Master's Thesis, University of California] Berkeley : University of California Press,

Thompson, R. K., & Pritchard, A. W. (1969). Respiratory adaptations of two burrowing crustaceans, *Callianassa californiensis* and *Upogebia pugettensis* (Decapoda, Thalassinidea). The Biological Bulletin, 136(2), 274-287.

CHAPTER 1: MORPHOLOGICAL COMPARISON OF OFFSHORE *NEOTRYPAEA* SP. WITH THE ESTUARINE *N. GIGAS* AND *N. CALIFORNIENSIS*

1. Introduction

The burrowing shrimp genus *Neotrypaea* originally belonged to the *Callianassa* genus of mud shrimp until Manning & Felder (1991) proposed to separate the two, and there has been ongoing discussion of the monophyletic status of *Neotrypaea* (Tudge et al. 2000, Sakai 2005, Campos et al. 2009). Within the *Neotrypaea* genus, there are seventeen distinct species including three found in west coast estuaries: *Neotrypaea biffari* (Holthuis, 1991), *N. gigas* (Dana, 1852*),*and *N. californiensis* (Dana, 1852). *N. biffari* or the 'tidepool ghost shrimp' burrows in sand-gravel substrate in the rocky intertidal and occurs from Goleta, California, to Tortugas Bay, Mexico (Campos et al 2009). This tidepool ghost shrimp is morphologically characterized by triangular eyes, with eyestalk tips undiverging (Campos et al., 2009). *Neotrypaea californiensis* (Dana, 1854) and *N. gigas* (Dana, 1852) are found in the estuaries and intertidal habitats of Oregon (Griffis and Chavez, 1988, Campos et al., 2009).

While *N. gigas* and *N. californiensis* (Figure 1) coexist in the estuaries in the west coast of the United States, they have distinct morphologies, distributions, and burrow types. *N. californiensis* are predominant in estuaries and exist in the sandy, upper intertidal substrate whereas *N. gigas* are less commonly found and occur in the lower to mid-intertidal sandy-mud substrate (Griffis and Chavez 1988). The morphological characters distinguishing the two northern Pacific estuarine species include distinctions in the eyestalk shape and the male major claw size and shape. *N. gigas* can be distinguished by the lack of gap and the straight edge of the dorsal ridge of the male major claw (Hart 1982, Sakai 1999, Pernet et al. 2010) as well as the outer distal edge of the eyestalk and the length of the eyestalk greater than the second article (Pernet et al. 2010). Additional features distinguishing *N. gigas* and *N. californiensis* according to Pernet et al. (2010) included the chela length being greater than the carpus length for the male major claw, the length of the eyestalk extending past the second article of the first antennae, and the concave shape of

the eyestalk. They are both detritivores, and their large claw is used either for mating behavior or for combat purposes rather than consumption (Felder and Lovett, 1989; Rowden and Jones 1994).

In 2019, a new population of an unknown *Neotrypaea* species was discovered 12 km offshore Newport, OR, with burrow opening densities of up to 400 per m^{-2} (Henkel *et al*. 2022). The offshore population of *Neotrypaea* more closely resembled the estuarine species of burrowing shrimp, *Neotrypaea gigas*, based on the large male claw and the lateral edges of the claw carpus and propodus (Pernet et al., 2010). Preliminary genetic testing using the 16S (metazoan short fragment) on five offshore shrimp was conducted by the Baker Lab at OSU. The 16S rRNA sequences equally matched to *N. gigas, N. californiensis, N. japonica, N. caesari, N. uncinata, N. tabogensis,* and *N. harmondii.* Many of those species were outside of the probable range of origin for the offshore population in Newport, OR. *N. uncinata* is located in the southeast Pacific from the southern tip of Mexico to the Peninsula ode Tiatao; Chile (Thatje, 2003), and *N. caesari* is found in Trinidad in the Lesser Antilles (Heard & Manning, 2000). The objective of this chapter was to conduct morphological analysis of individuals from the newly-discovered offshore population and compare to individuals of species of *Neotrypaea* resident in a nearby estuary in order to identify the offshore *Neotrypaea* sp. to the species level and determine any unique physical markers.

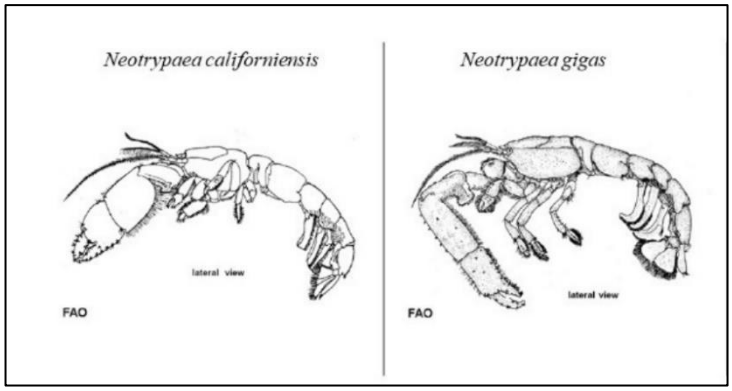


Figure 1. FAO Illustrations of N. californiensis, N. gigas. (Holthius, L.B. 1991; Hart, 1982; Liu, 1955).

In addition to identifying the shrimp to the species level, morphological distinctions between either species could offer insight into any environmental changes or functional roles of the offshore population. Mouth parts (maxillipeds) of the *Callianassa subterranea* have been identified as having specific function for

detritus feeding and sediment filtering (Stamhuis et al. 1998). Functional morphological changes may have resulted from being in a new environment and these functional advantages could have helped *Neotrypaea* sp. successfully persist in the offshore bed.

Figure 2. Female (left) and Male (right) Neotrypaea sp. collected at the PacWave South Site. Scale Bars are in centimeters.

Hypothesis 1: There is no difference in the morphological measurements between offshore *Neotrypaea* sp. and those in the estuary *Neotrypaea gigas* because they are the same species.

Hypothesis 2: The measurements of the estuarine *N. californiensis* will be distinct from those of the offshore *Neotrypaea* sp. as well as *N. gigas*, including the angle of the outer distal edge of the eyestalk and the gap between the propodus and dactyl of the male major claw.

2. Methods

Sample Collection

In 2019 - 2021, offshore *Neotrypaea* were discovered and collected using Grey-O'Hara 0.1 m^2 box cores offshore Newport, OR, USA. *N. gigas* and *N. californiensis* were collected from Yaquina Bay, Oregon, and Gray's Harbor, Washington, in 2021. The samples in Yaquina Bay were collected by members of the Henkel and Dumbauld labs, and those from Gray's Harbor were collected by Dr. Brett Dumbauld. All samples were stored in 70% EtOH until analysis.

Evaluation of Morphological Characters

Calipers were used to measure carapace length of the shrimp, and photos of the whole *Neotrypaea* next to a ruler were taken to measure appendages that exceeded the size of the microscope. Photographs of the major body parts identified by Pernet et al. (2010) as being potentially distinguishing between species were taken. Photos of the minor claw, walking legs, major claw the rostrum were taken under the compound microscope with Leica software and measurements were recorded using ImageJ.

Measurements were taken of 31 offshore *Neotrypaea* sp., 31 estuarine *N. gigas*, and 24 estuarine *N. californiensis.* These measurements included the lengths and heights of common parts (pollex, dactyl, propodus, carpus, merus, eyestalk); rostrum angle and gap between the chela and pollex of the major claw were included. Unique measurements were made as well including the cleft extending from the propodus to the chela, eyestalk gap, cornea width and area, and the carpus angle of the first walking leg. In total, 33 characteristics were evaluated. Some of the *Neotrypaea* parts were lost upon collection, which resulted in a lower sample size for most measurements (eyestalk, walking leg 01, walking leg 02, minor claw, major claw); two *Neotrypaea* had damaged corneas

Statistical Analysis

These measurements were all standardized to the carapace length. Q-Q plots were used to test for normality of data, and Levene's test for equal variance was conducted on the measurements. One-way ANOVA test and Tukey post-hoc tests were performed on each of the characteristics to determine statistically significant differences between each pair of populations. Linear Discriminant Analysis (10% withheld testing data) was performed on 26 characters including the eyestalk width, cornea.eyestalk width, cornea area.eyestalkwidth.carapacelength, eyestalk length, crest angle, eyestalk gap, all walking leg 01 characters and minor claw characters. Individuals and characters were withheld from the LDA to reduce the number of missing data, or NAs, in the matrix. Firstly, individuals were deleted if they were

missing their appendage measurements (such as an individual missing walking legs and major claw). Characters, or columns, were deleted if they contained more than 2 NAs.

3. Results

There were many distinct characteristics between the estuary *N. californiensis*, the offshore *Neotrypaea* and estuary *N. gigas:* 8/33 characteristics were significantly different between the *N. californiensis* and the offshore shrimp whereas 8/33 differentiated *N. californiensis* and the estuary *N. gigas*. Three characteristics were statistically significant in differentiating the estuary *N. gigas* and the offshore *Neotrypaea sp.* (Tables 1-5). The *N. californiensis* measured had carapace lengths between 7.2 and 17.02 mm, *N. gigas* had carapace lengths between 5.47 and 37.31 mm, and the offshore shrimp between 5.65 and 25.07 mm.

Eyestalks

The angle of the outer distal edge of the eyestalk was not significantly different between the offshore *Neotrypaea* and the estuary *N. gigas* (Table 1) while both differed from estuarine *N. californiensis*. The corneas of the offshore population were noticeably large, and the differences among the three populations was confirmed by the measurements of cornea width/eyestalk width/CL (one-way ANOVA, p-value < 0.001). The mean cornea width to eyestalk width ratio for the estuarine *N. gigas* was 0.253 whereas the mean cornea area of the offshore shrimp was 0.417 and *N. californiensis* was 0.428 (Table 1, Figure 3, Figure 4).

Walking Legs 01 and 02

None of the walking leg 01 or 02 characteristics were distinct among the three populations (Figures 5 & 6, Tables 2&3).

Minor Claw

There were no statistically significant differences between the offshore and the estuarine *N. gigas* measurements in the minor claw (Figure 7, Table 4). However, 4/7 characteristics were distinct between *N. californiensis* and the offshore population. Carpus height, carpus length, propodus height, and merus height were all distinct between *N. gigas* and *N. californiensis* (Figure 7, Table 4).

Major Claw

None of the male major claw characteristics were distinct between the offshore and estuary *N. gigas* (Figure 8, Table 5). The Propodus H/Propodus L was significantly different between *N. gigas* and *N. californiensis as* well as *N. californiensis* and the offshore population (Table 5).. Although not statistically significant, there were some gaps between the dactyl and propodus or the "fixed finger" of the male major claw in the offshore population, but none of the estuary *N. gigas* had a gap.

Overall

Based on the eyestalk, walking leg 01, and minor claw characters (which had the largest sample sizes), the three populations formed distinct clusters in the linear discriminant analysis and the model accuracy for the LDA was 0.75 (Figure 9).

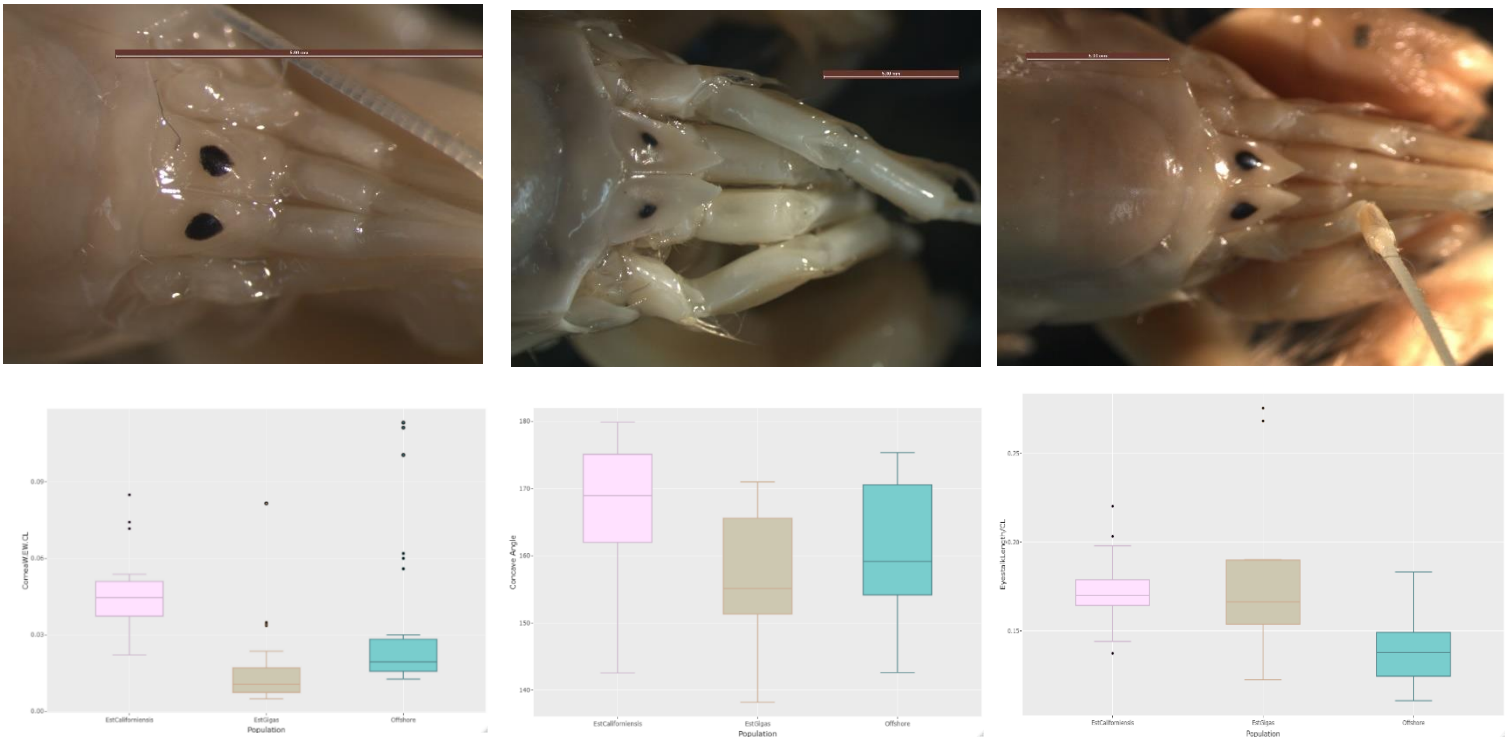


Figure 3*. Eyestalks of N. californiensis (first/pink) N. gigas (middle/beige) and offshore Neotrypaea (right/blue) with box plots of CorneaW/EW/CL, Concave Angle, and Eyestalk Length/CL.*

Table 1. *Averages, ANOVA p-value, and Tukey Post-Hoc p-values for eyestalk characteristics of N. gigas, Offshore, and N. californiensis. Significant results are bolded.*

	N. gigas	**Offshore**	***N. cal***	**ANOVA**	**Tukey Post-Hoc (p-value)**		
Eyestalk	n = 12 avg	n = 18 avg	n=24 avg	p-value	Offshore-EstGigas	EstCal-EstGigas	Offshore -EstCal
Eyestalk Length/CL	0.1794	0.1394	0.1718	**0.000**	**0.001**	0.720	**0.001**
Eyestalk Width/CL	0.0943	0.0897	0.1027	0.171			
Concave Angle	157.14	159.976	168.08	**0.007**	0.740	**0.012**	**0.041**
Eyestalk Gap/CL	0.0453	0.0512	0.0395	**0.018**	0.452	0.432	**0.013**
Cornea Area/Eyestalk Width/CL	0.0087	0.0156	0.0187	**0.000**	**0.010**	**0.000**	0.372
CorneaW/EyestalkWidth/ CL	0.0159	0.0313	0.0459	**0.000**	**0.018**	**0.000**	0.068

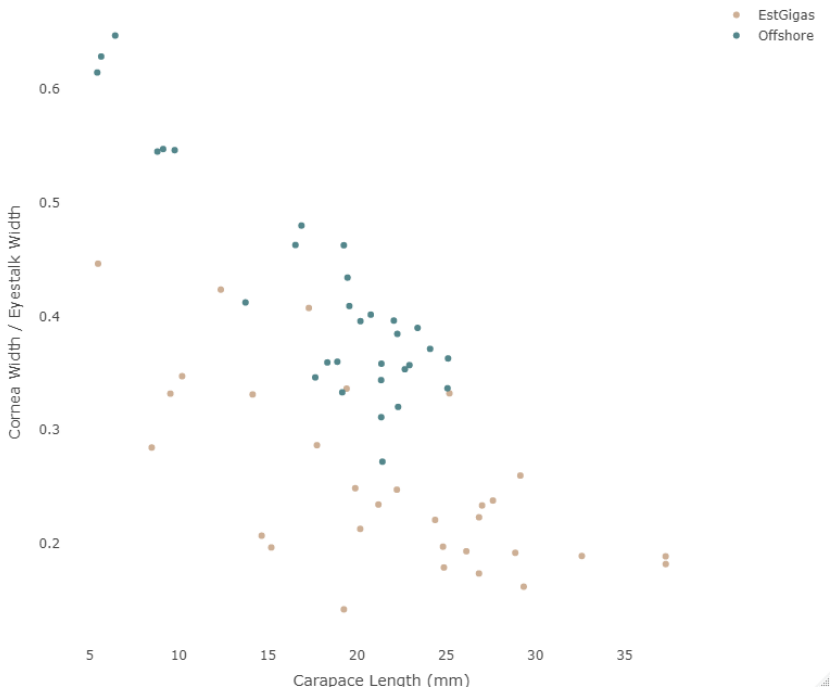


Figure 4*. Cornea Width/Eyestalk Width vs Carapace Length of Offshore (blue) Neotrypaea and Estuarine (brown) Neotrypaea gigas. Average for N. gigas = 0.253, Average for Offshore = 0.417.*

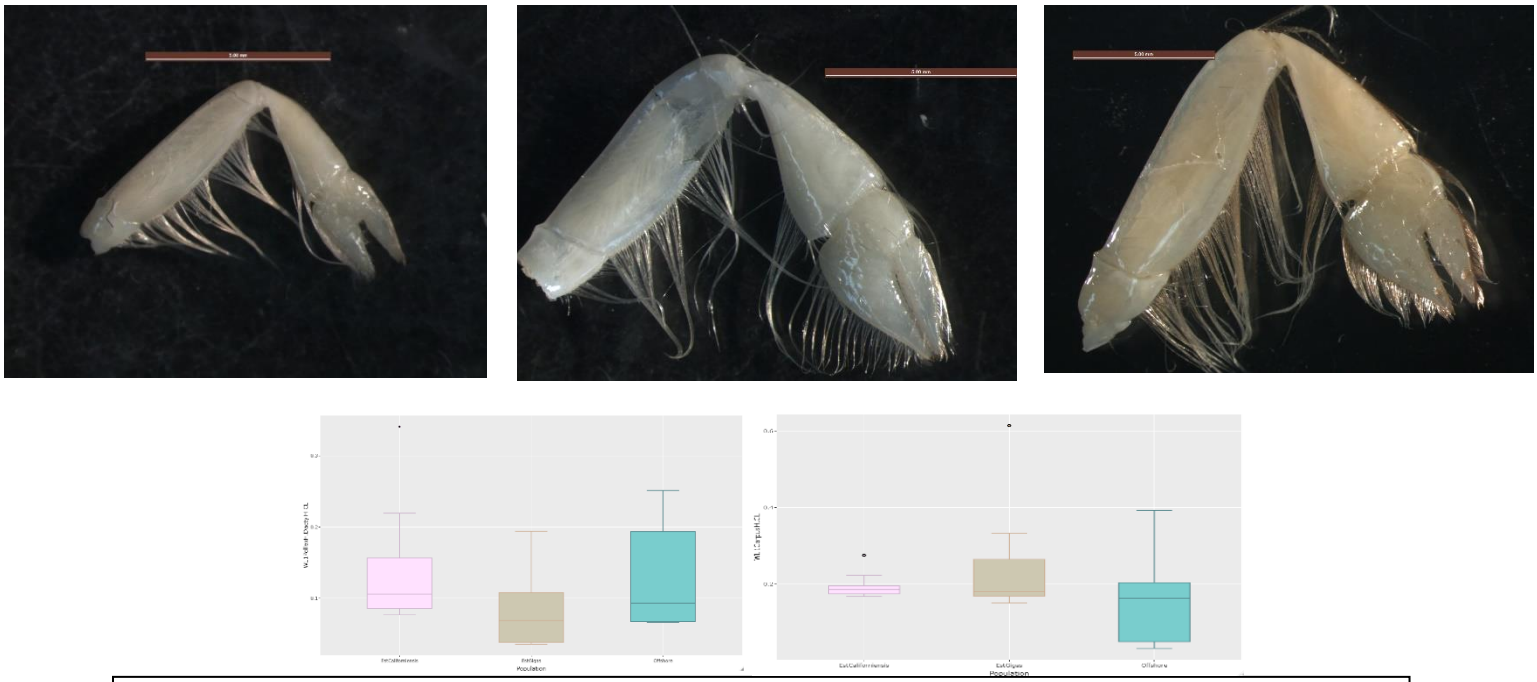


Figure 5*. Walking Leg 01 of N. californiensis (first/pink) N. gigas (middle/beige) and offshore Neotrypaea (right/blue) with box plots of the pollexh/dactylh ratio and carpus height both standardized by carapace length.*

Table 2. *Averages and ANOVA p-values Walking Leg 01 characteristics of N. gigas, Offshore, and N. californiensis. No significant differences were found among populations.*

	N. gigas	**Offshore**	***N. cal***	**ANOVA**
Walking Leg 01	n=11 avg	n=17 avg	n=14 avg	p-value
PropodusL/CL	0.3269	0.3271	0.3066	0.860
Gap/CL	0.0447	0.0414	0.0571	0.803
CarpusH/CL	0.2403	0.2173	0.1924	0.341
PollexH/DactylH/CarapaceL	0.0783	0.0952	0.1363	0.078
CarpusA	44.207	47.119	44.681	0.579

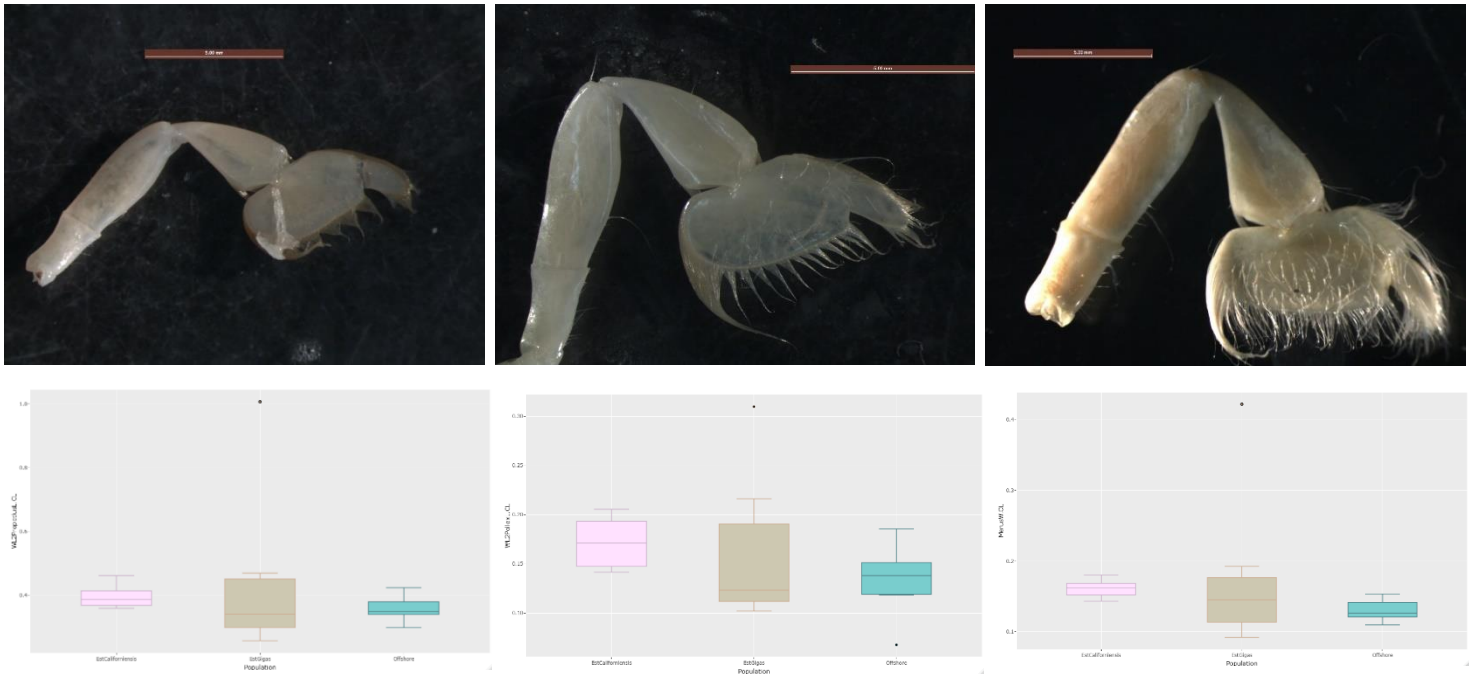


Figure 6*. Walking Leg 02 of N. californiensis (first/pink) N. gigas (middle/beige) and offshore Neotrypaea (right/blue) with box plots of Propodus Length/CL, Pollex Height/CL, and MerusWidth/CL.*

Table 3. *Averages and ANOVA p-values for Walking Leg 02 characteristics of N. gigas, Offshore, and N. californiensis. No significant differences were found among populations.*

	N. gigas	**Offshore**	***N. cal***	**ANOVA**
Walking Leg 02	n=9 avg	n=12 avg	n=11 avg	p-value
PropodusL/CL	0.4222	0.3569	0.3937	0.489
PropodusH/CL	0.2053	0.1775	0.2096	0.397
PollexL/ CL	0.155	0.1360	0.1714	0.161
PollexW/CL	0.0687	0.0646	0.0726	0.638
CarpusW/CL	0.214	0.1781	0.1956	0.478
CarpusL/CL	0.3439	0.3239	0.3689	0.445
MerusW/CL	0.154	0.1299	0.1607	0.188

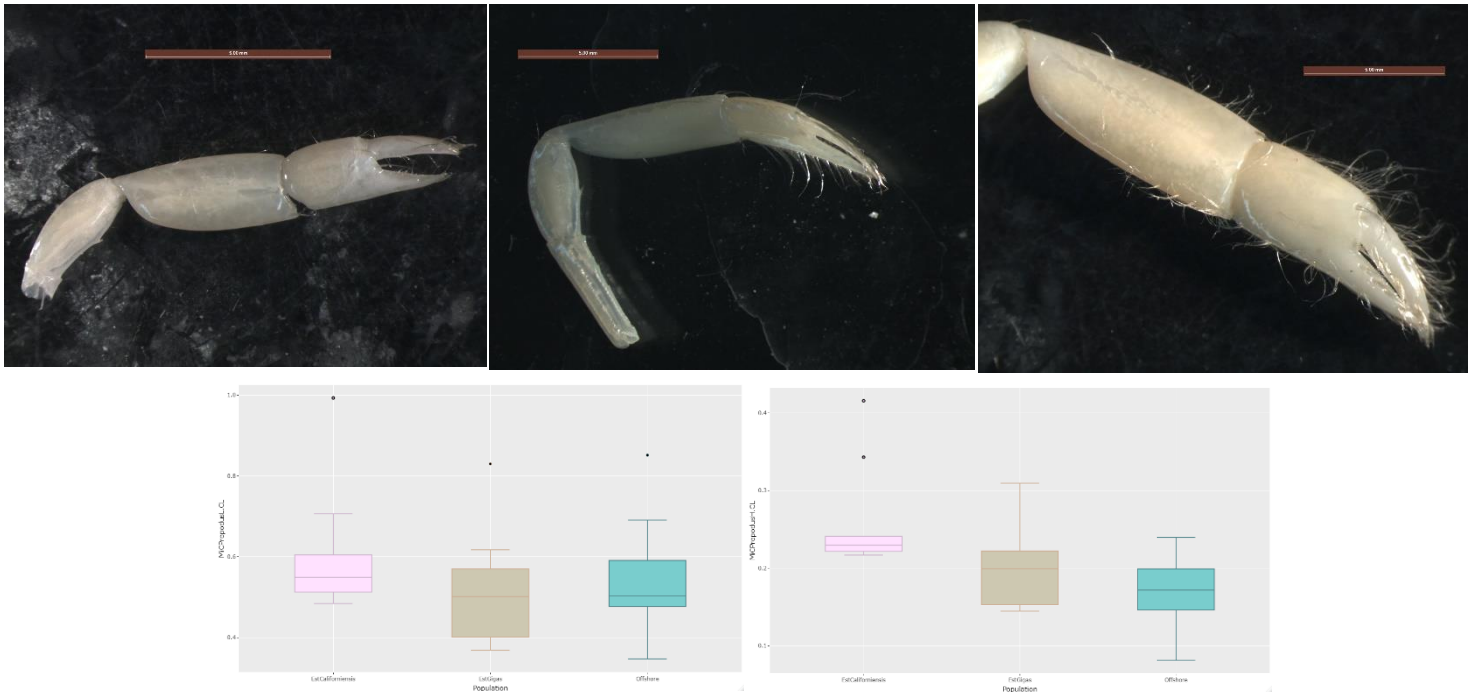


Figure 7*. Minor claw of N. californiensis (first/pink) N. gigas (middle/beige) and offshore Neotrypaea (right/blue) with box plots of propodus length and propodus height standardized to carapace length.*

Table 4. *Averages, ANOVA p-values, and Tukey post-hoc p-values for the Minor Claw characteristics of N. gigas, Offshore, and N. californiensis Significant results are bolded.*

	N. gigas	**Offshore**	***N. cal***	**ANOVA**	**Tukey Post-Hoc (p-value)**		
Minor Claw	n=11 avg	n=13 avg	n=10 avg	p-value	Offshore - EstGigas	EstCal-EstGigas	Offshore-EstCal
PropodusL/CL	0.5088	0.5426	0.6021	0.305			
PropodusH/CL	0.1958	0.1714	0.2579	**0.002**	0.499	**0.028**	**0.001**
CarpusH/CL	0.1674	0.1814	0.2817	**0.003**	0.897	**0.004**	**0.011**
CarpusL/CL	0.4441	0.5753	0.6953	**0.025**	0.268	**0.019**	0.348
DactylH/CL	0.0929	0.0847	0.1136	**0.031**	0.702	0.159	**0.026**
PollexH/CL	0.0827	0.0738	0.1032	**0.003**	0.494	0.050	**0.003**
MerusH/CL	0.1038	0.1724	0.2367	**0.002**	0.102	**0.001**	0.131

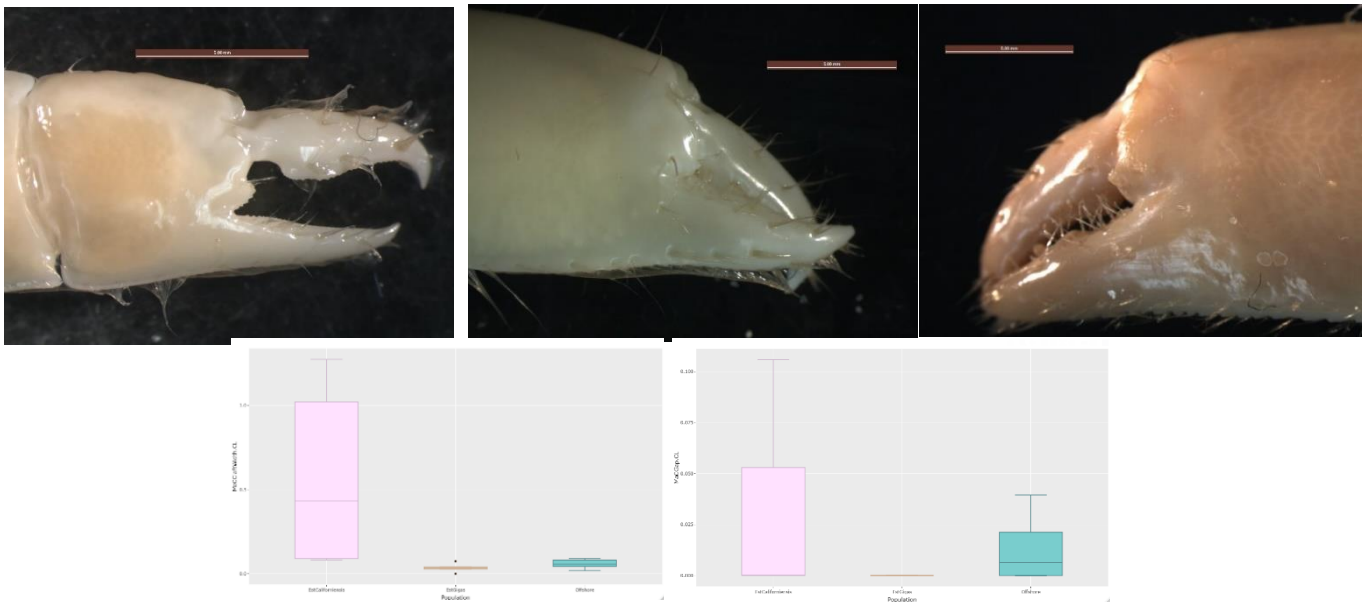


Figure 8*. Male Major Claw of N. californiensis (first/pink) N. gigas (middle/beige) and offshore Neotrypaea (right/blue) with box plots of cleftwidth/CL, dactyl-propodus gap/CL.*

Table 5. *Averages, ANOVA p-values, and Tukey post-hoc p-values for the Major Claw characteristics of N. gigas, Offshore, and N. californiensis. Significant results are bolded.*

	N. gigas	**Offshore**	***N. cal***	**ANOVA**	**Tukey Post-Hoc (p-value)**		
Major Claw	n=6 avg	n=12 avg	n=4 avg	p-value	Offshore-EstGigas	EstCal-EstGigas	Offshore-EstCal
PropodusL/CL	1.0330	0.9594	0.8364	0.691			
PropodusL/PropodusH	1.8750	1.9132	1.5397	**0.000**	0.913	**0.031**	**0.007**
CleftWidth/CL	0.0354	0.0586	0.0610	0.184			
PropodusH/CL	0.5387	0.5049	0.5623	0.832			
CarpusL/CL	0.6395	0.7183	0.6320	0.685			
CarpusH/CL	0.5113	0.4740	0.5708	0.607			
PollexH/CL	0.1964	0.1817	0.1685	0.785			
DactylH/CL	0.2129	0.2240	0.2036	0.920			
Dactyl-Propodus Gap/CL	0.0000	0.0115	0.0265	0.241			

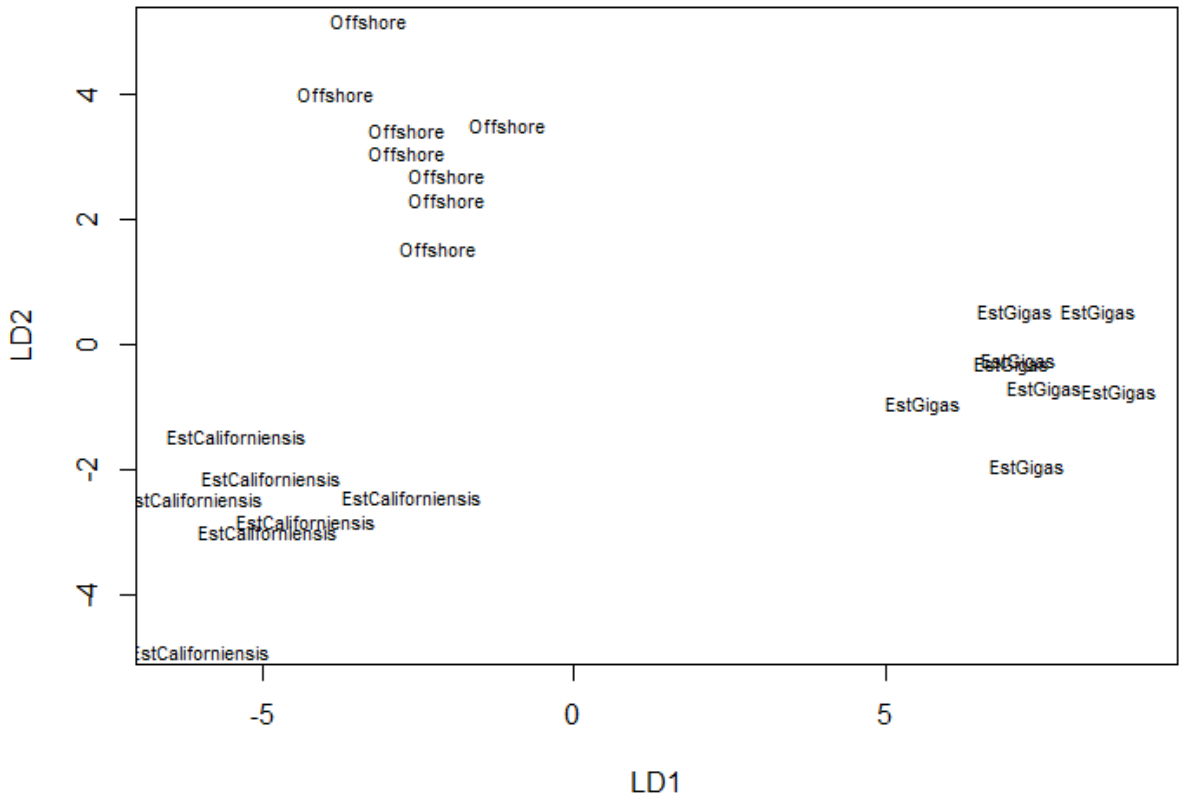


Figure 9. *Linear Discriminant Analysis (LDA) Plot of the three populations (Estuary N. californiensis, Estuary N. gigas, and Offshore Neotrypaea sp.). 26 Characters included measurements from the eyestalks, minor claw, and walking leg 01. Testing and taining data were split up into 90% and 10%. Model Accuracy was 0.75.*

4. Discussion

After testing for the characteristics first identified by Pernet et al. (2010), the eyestalk curvature was the only consistent characteristic that differentiated both *N. gigas* from *N. californiensis* and the Offshore from *N. californiensis* while remaining the same between *N. gigas* and the Offshore. The curved eyestalks of the *N. gigas* and *N. californiensis* were the primary characteristic used by Pernet et al. This was our initial observation, and it was confirmed

by the measurement of the concave angle. Given the lack of male major claws in our sample sizes, eyestalk characters should continue to be used for identifying the two species of *Neotrypaea*. Although the dactyl-propodus gap of the major claw remained useful in differentiating *N. californiensis* from *N. gigas* in the estuary, there were small gaps found in the male major claw of five of the 12 offshore *Neotrypaea* sp., which were never found in the estuary *N. gigas*. A greater sample size for the male major claw, of which the estuary *N. gigas* only had 6, may have resulted in a statistically significant ANOVA result for this feature.

Based on the three distinct measurements found between the estuary *N. gigas* and the offshore *Neotrypaea*, we fail to reject the original hypothesis that there would be no distinctions between *N. gigas* and the new population. This was confirmed by both the LDA analysis and the measurements of the eyestalk length, cornea area.eyestalkwidth.carapacelength, and cornea width.eyestalkwidth.carapacelength, The most apparent trait difference between the estuary *N. gigas* and the offshore *Neotrypaea* was the width of the cornea which was noticed as larger in the offshore shrimp than the estuarine *N. gigas* while making other measurements and subsequently analyzed. A similar disparity was noted in the cornea width/eyestalk width of two species of *Neotrypaea*: *N. japonica* and *N. harmandii* endemic to Japan (Manning and Tamaki 1998, Wardiatno and Tamaki 2001). These were originally named *Nihonotrypaea*, but this has been debated with *Nihonotrypaea* and *Neotrypaea* being clustered together in a monophyletic clade with subclades not distinguishing *Nihonotrypaea* (Felder and Robles 2009) but current literature still refers to *Neohonotrypaea* (Tamaki et al 2020). The larger cornea size of the offshore as compared to the estuarine *N. gigas* does resemble the difference depicted by Wardiatno and Tamaki 2001. *N. harmandi* (cornea width/eyestalk width ratio > 0.5) is typically found in the intermediate waters of Tamioka Bay in Japan, whereas *N. japonica* (cornea width/eyestalk width

ratio < 0.5) is found in the estuary Ariake Sound in Kyushu, Japan (Wardiatno and Tamaki 2001). For comparison, the offshore *Neotrypaea* had an average cornea W: EW ratio of 0.41 and the estuarine *N. gigas* had an average of 0.25 not accounting for CL. The estuarine *N. californiensis* had an average cornea W: EW ratio ratio of 0.43 which was similar to that of the offshore *Neotrypaea*.

Differentiations between the offshore population and *N. gigas* should be pursued through genetic analysis. Although preliminary genetic testing has already been conducted, more samples should be submitted with specific markers to identify the offshore population. The COI (cytochrome oxidase subunit 1) marker was prepped as a working reference for *N. gigas* with the preliminary genetic samples, but samples have not yet been run. This marker was used by Pernet et al. (2008) and is commonly used in genetic research of *Neotrypaea* (Buncic 2010, Gille 2012). Further genetic testing could point to a hybrid species of *N. californiensis* and *N. gigas* or at least genetic variation across location and habitat spread. More morphological analysis of *Neotrypaea* collected offshore at locations outside of the PacWave South site and at multiple estuaries along the west coast would offer an in depth perspective on morphological variation. The morphological differences found in this analysis – in the cornea and eyestalks - bring to question the true origin of the offshore population.

The morphological distinctions found in the new offshore *Neotrypaea* sp. have broader implications for who exactly they are. Although not statistically significant, a gap was observed in the offshore population of the male major claw. The male major claw in *Neotrypaea* is distinctive based on maturity of the individual and can occur on either the left or the right side, and that of *N. californiensis* is thought to be used in grappling (Labadie and Palmer, 1996). The

gape would make it difficult to grasp other smaller limbs, and the greater size of the propodal notch increases the likelihood of the fixed finger snapping under pressure (Labadie and Palmer 1996). This was reflected too in some male major claws of *N. californiensis* being damaged, whereas none of the *N. gigas* male major claws were damaged. There may be some indication that this offshore population could have a functional advantage, or disadvantage, offshore. Changes in the morphology of the claws may have resulted in a competitive disadvantage returning to the estuary, or the cornea size could be an advantage in deeper waters with low visibility. The large corneas (similar to *N. californiensis*) and the presence of a small gape in some male major claws could indicate a hybrid species between *N. californiensis* and *N. gigas* that has successfully persisted offshore.

The resulting possibilities for the identity of the offshore populations include (1) a morphologically distinct *N. gigas* (2) a hybrid between *N. gigas* and *N. californiensis* or (3) a unique offshore *Neotrypaea* species or (4) a species derived from elsewhere. Based on these morphological findings, the most likely possibility would be (1) or (2). If a morphologically distinct *N. gigas*, the characteristics could offer an advantage in deeper waters with large corneas in conditions of worse visibility. Or, they could have some morphological disadvantage that led them to fail to compete with the young *Neotrypaea* returning to shore.

5. References

Buncic, M. (2010). *Interannual differences in the estuarine ghost shrimp, Neotrypaea californiensis*. .[Masters Thesis, San Jose State University] San Jose State University ProQuest Dissertations Publishing

Dana, J. D. (1852). Macroura. Conspectus Crustaceorum & Conspectus of the Crustacea of the Exploring Expedition under Capt. C. Wilkes, U.S.N. Proceedings of the Academy of Natural Sciences of Philadelphia 6: 10-28.

de Campos, A., Campos, E., & Manriquez, I. (2009). Intertidal thalassinidean shrimps (Thalassinidea, Callianassidae and Upogebiidae) of the west coast of Baja California, Mexico: annotated checklist, key for identification, and symbionts. Crustaceana, 82(10), 1249-1263.

Columbia, B., & British Columbia Provincial Museum. (1982). Crabs and their relatives of British Columbia. British Columbia Provincial Museum

Felder, D. L., Robles, R. A. F. A. E. L., Martin, J. W., & Crandall, K. A. (2009). Molecular phylogeny of the family Callianassidae based on preliminary analyses of two mitochondrial genes. Decapod crustacean phylogenetics. Crustacean issues, 18, 327-342.

Gille, D. A. (2012). *Genetic population structure and cryptic speciation of ghost shrimp (Neotrypaea californiensis) in North American West Coast estuaries*. [Masters Thesis, San Jose State University] San Jose State University ProQuest Dissertations Publishing

Griffis, R. B., & Chavez, F. L. (1988). Effects of sediment type on burrows of *Callianassa californiensis* Dana and *C. gigas* Dana. Journal of Experimental Marine Biology and Ecology, 117(3), 239-253.

Heard, R. W., & Manning, R. B. (2000). A new genus and species of ghost shrimp from Tobago, West Indies (Crustacea: Decapoda: Callianassidae). Proceedings of the Biological Society of Washington.

Holthius, L. B. (1991). FAO Species catalogue–marine lobsters of the world: An annotated and illustrated catalogue of marine lobsters known to date. Volume, 13, 128-132.

Labadie, L. V., & Palmer, A. R. (1996). Pronounced heterochely in the ghost shrimp, Neotrypaea californiensis (Decapoda: Thalassinidea: Callianassidae): allometry, inferred function and development. *Journal of Zoology*, *240*(4), 659-675.

Manning, R. B., & Tamaki, A. (1998). A new genus of ghost shrimp from Japan (Crustacea: Decapoda: Callianassidae). Proceedings of the Biological Society of Washington.

Milne Edwards, H. (1837) Histoire naturelle des Crustacés, comprenant l'anatomie, la physiologie et la classification de ces animaux. Vol. 2., Librairie Encyclopédique de Roret, Paris.

Pernet, B., Deconinck, A., & Haney, L. (2010). Molecular and morphological markers for distinguishing the sympatric intertidal ghost shrimp *Neotrypaea californiensis* and *N. gigas* in the Eastern Pacific. Journal of Crustacean Biology, 30(2), 323-331.

Sakai, K. (1999). Synopsis of the family Callianassidae, with keys to subfamilies, genera and species, and the description of new taxa (Crustacea: Decapoda: Thalassinidea). Zoologische Verhandelingen, 326, 1-152.

Sakai, K. (2005). Callianassoidea of the World:(Decapoda, Thalassinidea). Brill.

Somiya, R., Suzuki, T., & Tamaki, A. (2014). Mouthpart morphology and wild diet of zoeae of the ghost shrimp, *Nihonotrypaea harmandi* (Decapoda: Axiidea: Callianassidae). Journal of Crustacean Biology, 34(3), 300-308.

Stamhuis, E. J., Dauwe, B., & Videler, J. J. (1998). How to bite the dust: morphology, motion pattern and function of the feeding appendages of the deposit-feeding thalassinid shrimp *Callianassa subterranea*. Marine Biology, 132(1), 43-58.

Tamaki, M., Wang, Z., Barnes-Diana, T., Guo, D., Berard, A. V., Walsh, E., ... & Sasaki, Y. (2020). Complementary contributions of non-REM and REM sleep to visual learning. Nature neuroscience, 23(9), 1150-1156.

Thatje, S. (2003). Una Revisión de los Thalassinidea (Crustacea: Decapoda) de Chile y Argentina. Review of the Thalassinidea (Crustacea: Decapoda) from Chile and Argentina. In Anales del Instituto de la Patagonia (Vol. 31, pp. 115-122).

Tudge, C. C., Poore, G. C., & Lemaitre, R. (2000). Preliminary phylogenetic analysis of generic relationships within the Callianassidae and Ctenochelidae (Decapoda: Thalassinidea: Callianassoidea). Journal of Crustacean Biology, 20(5), 129-149.

Wardiatno, Y., & Tamaki, A. (2001). Bivariate discriminant analysis for the identification of *Nihonotrypaea japonica* and *N. harmandi* (Decapoda: Thalassinidea: Callianassidae). Journal of Crustacean Biology, 21(4), 1042-1048.

CHAPTER 2: SETTLEMENT AND SULFATE-REDUCING BACTERIA ASSOCIATED WITH THE SEDIMENT OF OFFSHORE AND ESTUARINE *NEOTRYPAEA* SP.

Introduction

In pacific estuarine tidal flats, *Neotrypaea* sp. can occupy >80% of the intertidal area of some estuaries (Leduc and Pilditch 2017, DeWitt 2004); burrows increase the total surface area of sediment exposed to water and available for microbial colonization, including those responsible for ammonia and sulfide oxidation (MacGinitie 1934, Aller and Aller 1986, Griffis & Suchanek 1991). The microbial composition of *Neotrypaea* burrows and those of the surface sediments are highly correlated when they have a similar geochemical characteristic– ferric iron (Bertics and Zeibis 2009). The stage of settlement for benthic decapod crustaceans is unknown, with decapodid stage primarily thought of as the settlement stage but late zoeal stages also studied (Jensen 1991, Strasser and Felder 1999). Microbes are connected to the environmental characteristics of *Neotrypaea* habitat, as well as settlement of invertebrates to the benthos, and microbial sequencing of the sediment they inhabit would be a critical component to understanding their influence on the sediment as well as microbial influence on habitat suitability

Larval settlement, which occurs in June - November for *Neotrypaea californiensis* (Dumbauld *et al.* 1996), probably occurred in 2014-2015 for the offshore population based on cohort analysis (Henkel *et al*. 2022). Trends in estuaries over time have shown fluctuations in the recruitment and population size of *Neotrypaea*. Settlement of *N. californiensis* in Yaquina Bay, Oregon, increased from 151 shrimp m^{-2} in 2015 to 280 shrimp m^{-2} in 2016 after a sharp continuous decline since 2004 (Dumbauld and Bosley 2018). Therefore, the settlement of the larvae are related to the population size of *Neotrypaea* and their spatial distribution. Variations in the different environmental conditions and/or changes in microbes of the seafloor could result in a shift in the population size or distribution of *Neotrypaea*.

Microbes have the capacity to cue the settlement and metamorphosis of invertebrate larvae (Hadfield 1984, Dobretsov et al. 2020, Valesco et al. 2021). The settlement cues of ghost shrimp larvae are unknown and are critical to understanding what may induce their settlement offshore. There is variable research on substrate selection, organic and inorganic cues, and adult cues across species of ghost shrimp. *Neotrypaea californiensis* larvae have been shown to select for muddy areas rather than epibenthic shell cover for settlement in both field and laboratory experiments (Feldmen et al. 1997). A known settlement cue of burrowing shrimp *Callichirus major* and *Callichirus islagrande* includes a substrate (sand) independent of organic material or exposure to adults. However, burrowing activity at the decapodid stage does increase with exposure to natural sand/organic compounds (Strasser and Felder 1999**)**. In that same study, autoclaved sand was used to first expose the zoea stage before exposing decapodid stage to natural sand, indicating to the researchers that bacteria are not necessary to cue burrowing settlement of the decapodids. In another study, the substrate itself (sand) triggered settlement in the zoea of *C. major* from the Gulf of Mexico, but the Atlantic population of *C. major* required both the sand and water previously exposed to adults to trigger settlement (Strasser and Felder 1999). In estuarine habitat where adult populations were removed in support of oyster aquaculture, *N. californiensis* continue to recruit to the same location (Dumbauld and Bosley 2018). More research is necessary to fully determine the causes and pathway to settlement for *Neotrypaea*.

Microbes have been shown to trigger the settlement and metamorphosis of other invertebrates. *Bacillus* and *Cellulophaga* can the settlement of the tubeworm *Hydroides elegens,* (Huang and Hadfield 2003, Feckleton et al. 2017). *Shewanella* are known to cue the settlement of mussel (*Mytilus coruscus*) and oyster (*Crassostrea gigas*) larvae (Yang et al. 2013, Rischer *et*

al. 2016, Lebare and Weiner 1990). Another genus of interest to explore is *Pseudoalteromonas*, which is associated with the settlement of benthic invertebrates including tube worms (Hadfield 2011, Alker et al. 2020, Peng et al. 2020). In addition, it has been shown that the effects of temperature and concentration can influence the success of sponge larvae settlement – highest percent successful settlement on biofilms occurred at higher temperatures and longer biofilm development periods (Whalan and Webster 2014).

This chapter will focus on bacteria found in the sediments with *Neotrypaea* spp. as compared to sediments without *Neotrypaea* spp. in both offshore and estuarine habitats, which was conducted through sediment processing and sequencing. It will explore both the settlement aspect of specific microbes of the sediment, as well as the influence of *Neotrypaea* presence on the bacteria found in the sediments. Particularly, both sulfur-reducing bacteria which are found in anoxic sediments of the ocean, as well as cable bacteria which have been studied in the burrows of another burrowing shrimp, *Upogebia major*. Cable bacteria are the sole bacterium in the genus *Candidatus electrothrix*, which conducts electron transport to connect the sulfidic and oxic zones. They occurred in all the samples for the *Upogebia* shrimp burrow but not the sediment below the burrows – until that sediment was exposed to oxygen (Li et al. 2020). Microbial analysis of the sediments of *Neotrypaea* offers ecological insight into firstly why they occur where they do and secondly their role as microbial ecosystem engineers.

Hypothesis 1: There will be greater abundances of settlement-inducing bacteria (*Shewanella* sp, *Pseudoalteromonas* sp, *Bacillus* sp., or *Cellulophaga* sp. in sediment with *Neotrypaea* sp. both in the estuary and offshore.

Hypothesis 2: *Neotrypaea* presence in the sediment will induce changes in the microbial composition of the sediment, resulting in high similarity of taxa in samples containing *Neotrypaea* in both habitats.

2. Methods

2.1 *Sediment Sequencing*

Sediment was collected at four sites: Yaquina Bay, outside of the shrimp bed in an area with no evidence of burrowing activity (n=12) and within the *N. gigas* shrimp bed (n=12) and offshore from box cores taken from areas with (n=13) and without (n=11) *Neotrypaea* (Figure 10). These were then stored in the freezer (-4°C) until extraction. Extractions were then stored at (-40°C) before processing began as specified by kit protocol. A Quiagan DNeasy PowerSoil Pro Kit was used for the extraction of all samples of 250 mg of sediment from each of the four sites. Sequencing was performed at the Center for Quantitative Life Sciences at OSU, using the protocol described by Reimers et al. (2017). Primers for 16s amplification are: 357wF (5′-CCTACGGGNGGCWGCAG-3′) and 785R (5′-GACTACHVGGGTATCTAATCC-3′) through Illumina MiSeq.

Figure 10: *Locations of sediment samples collected, including sites in Yaquina Bay, and offshore Newport, Nehalem, and Nestucca Bay.*

These ran in multiple batches, the first batch of the estuary sediment in January on a 2x250bp (4 million reads/sample) and then another in May after the sediment was collected

offshore. The sediment samples collected offshore Nehalem were submitted to the CQLS during a back-log of samples due to a broken Illumina MiSeq Machine, so they were run on a loaner instrument from University of Oregon using Nano flow type (1 million reads/sample). The barcodes were also used and so the result of those final 8 samples were 100,000 reads/sample. Because of this, only the first 40/48 samples were used for ordination plots.

2.2 *Statistical Processing*

Sequences were analyzed using the DADA2 pipeline package for relative abundance at the genus level by comparing Amplicon Sequence Variants (ASVs). The DADA2 pipeline is a common and effective method of matching the sequences to a reference bacterial strain dataset (Callahan et al. 2016). This pipeline takes the Illumina-sequenced reads and tests the error in the number of reads per the number of unique sequences to standardize the sequences per sample. The data were normalized by converting the data from abundance to relative abundance by dividing reads of a specific ASV by the total number of ASVs. Multivariate plots with Bray-Curtis dissimilarity were used to visualize similarities or differences among groups (NMDS and PCA) and a PERMANOVA was used to determine the significance of location and *Neotrypaea* sp. presence on the centroids of the NMDS ordination plot. Abundances of *Desulfovibrionaceae* (containing cable bacteria), *Pseudoalteromonas* sp., *Shewanella* sp., *Bacillus* sp., and *Cellulophaga* sp. were individually compared among areas. Levene's Test was used to test for equal variance, and Shapiro-Wilk test was performed to determine normality of the data. A Two-Way ANOVA was used to determine the influence of Location and *Neotrypaea* sp. presence on the abundance of the microbial genera.

3. Results

3.1 Settlement Inducing Bacteria

The genera *Shewanella* sp. and *Pseudalteromonas* sp. were in higher relative abundance in sediment where *Neotrypaea* were present compared to the sediment without shrimp (Figure 11). An ANOVA test demonstrated that *Neotrypaea* sp. presence influenced relative abundance of *Shewanella* sp. but location did not (Table 6, $p < 0.001$, $p = 0.171$). *Pseudoalteromonas* sp. was influenced by location only (Table 6, $p = 0.003$). *Bacillus* sp. was found in the estuary but not offshore (Figure 11c). *Cellulophaga* sp. abundance differed significantly by location (Table 6, $p < 0.05$) but not the presence of *Neotrypaea* sp. (Figure 11d).

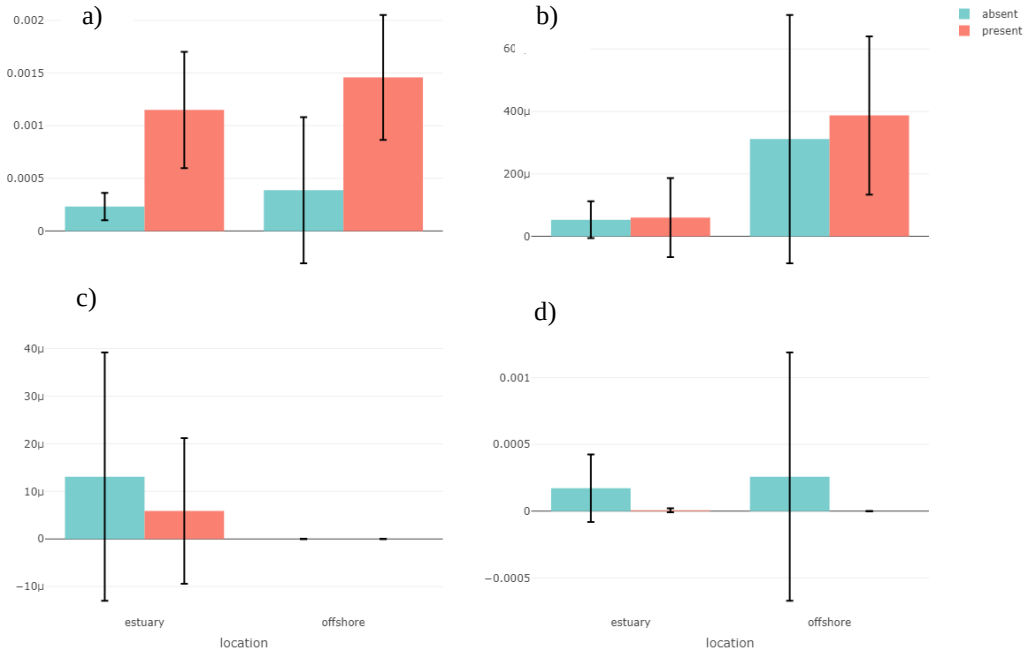


Figure 11. *Relative abundance of the four genera:* Shewanella (a), Pseudoalteromonas (b), Bacillus (c), Cellulophaga (d) *of interest for settlement, average of 48 samples.*

3.2 *Multivariate Community Patterns*

The NMDS using Bray-Curtis dissimilarity showed that the majority of the sediment samples collected share ASVs based on location and *Neotrypaea* sp. presence (Figure 12). NMDS1 separates out the offshore from the estuary locations and NMDS2 separates out the absences or presence samples. There is one *Neotrypaea* present sample that is clustered with the absences in the estuary. The centroids of the different sample sites are distinct based on a PERMANOVA ($p < 0.001$ for Location, *Neotrypaea*, and Location:*Neotrypaea*).

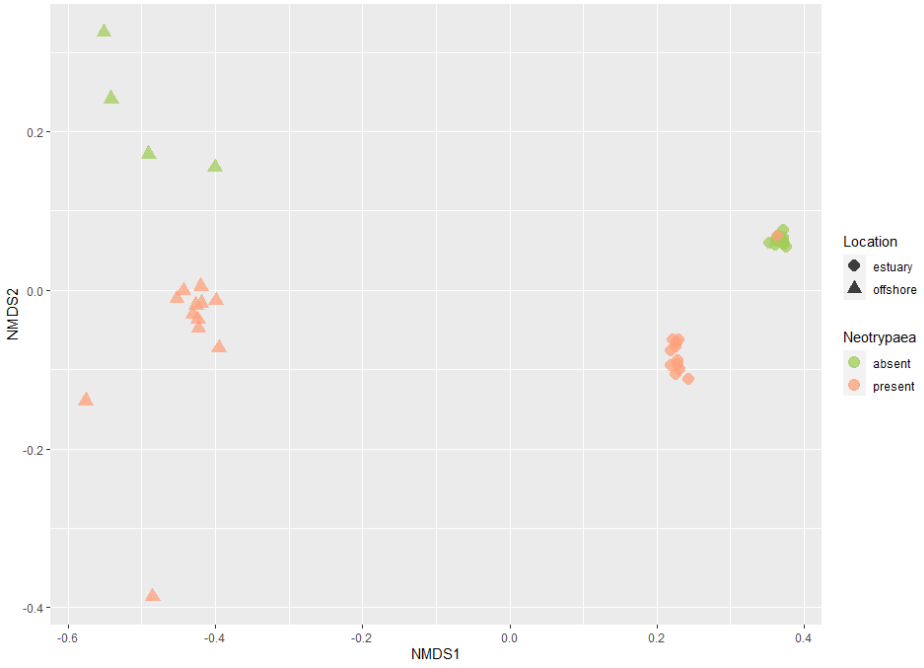


Figure 12. *NMDS plot using Bray-Curtis distance for the offshore and estuary sites (triangle/circle) and the Neotrypaea sp. presence or absence (green/salmon). Stress = 0.0258. Total Taxa = 141042.*

3.3 *Sulfate Reducers and Cable Bacteria*

There was a greater abundance of genera in the family *Desulfovibrionaceae* in the sediment with *Neotrypaea* sp. than without *Neotrypaea* sp. (Figure 13), although only *Desulfobolbus* was statistically significant based on the ANOVA (Table 6, $p < 0.0001$, $p = 0.171$). In addition to this family being more abundant in *Neotrypaea* sp. positive samples, a PCA plot revealed sulfate-reducing microbes distinguished the samples from others. The phyla *Desulfobacterota* and *Sva0485* distinguished the presence samples from the absence sediment (Figure 14).

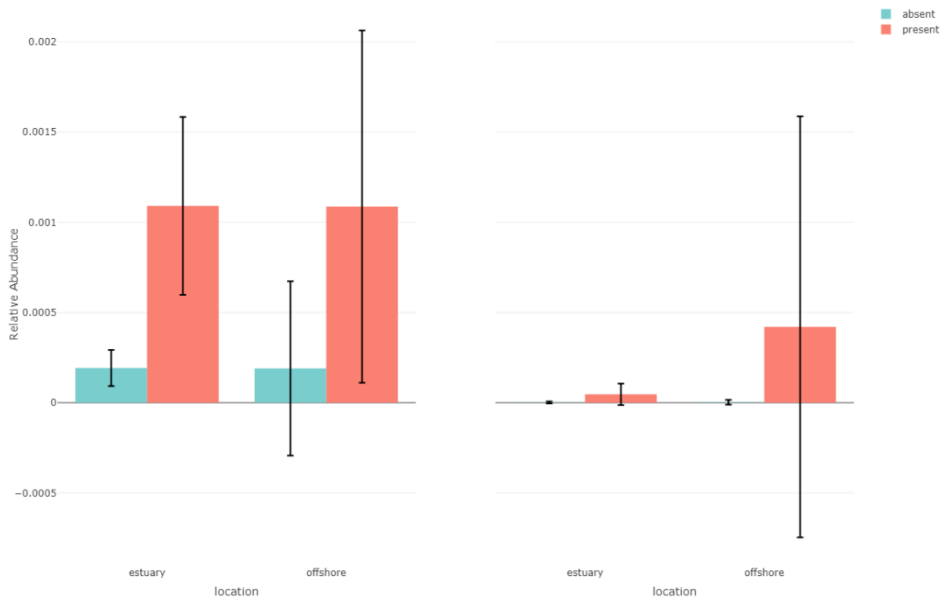


Figure 13. *Relative abundance the Genera Desulfobulbus* (Left) *and Desulfobulbus* (Right)*, in the family* Desulfovibrionaceae*. Values are averages of 48 samples with standard deviation.*

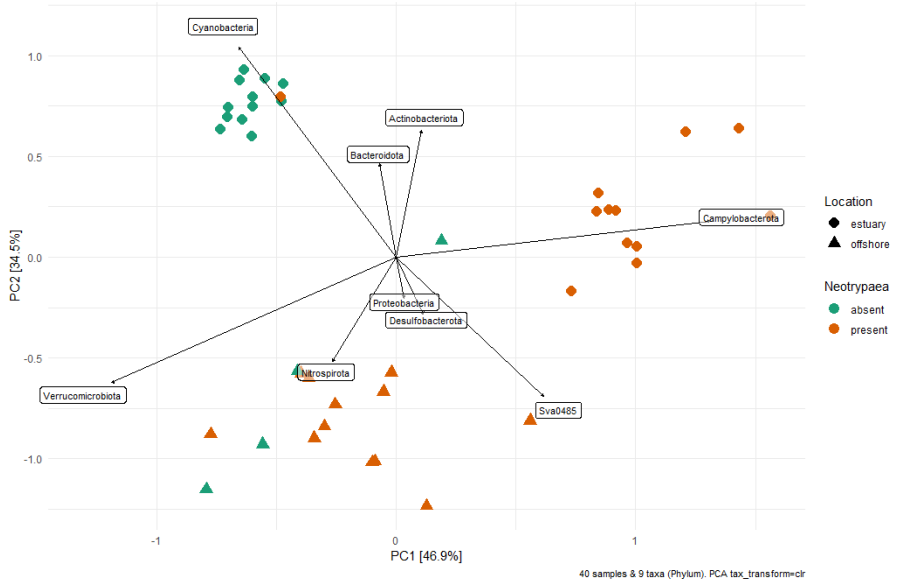


Figure 14. *PCA plot of 40 sediment samples categorized by Location - estuary or offshore (circle/triangle) and* Neotrypaea sp. *- presence or absence (salmon/green).*

Table 6. ANOVA results for each of the six genera of interest including those related to settlement.

	ANOVA		
Genera	**Factor**	**F-value**	**P value**
Shewanella	Neotrypaea	35.305	**4.11E-07**
	Location	1.937	0.171
	Neotrypaea:Location	0.209	0.65
Pseudoalteromonas	Neotrypaea	0.329	0.5694
	Location	9.799	**0.0031**
	Neotrypaea:Location	0.119	0.7323
Cellulophaga	Neotrypaea	0.841	0.3643
	Location	4.557	**0.0384**
	Neotrypaea:Location	0.679	0.4143
Bacillus	Neotrypaea	0.869	0.356
	Location	1.076	0.305
	Neotrypaea:Location	0.036	0.851
Desulfobulbus	Neotrypaea	26.64	**5.65E-06**
	Location	0	0.986
	Neotrypaea:Location	0	0.986
Electrothrix	Neotrypaea	1.976	0.167
	Location	1.318	0.257
	Neotrypaea:Location	1.202	0.279

4. Discussion

The main objectives of this chapter were to (1) identify bacteria known to cue invertebrate settlement that could possibly also induce *Neotrypaea* sp. settlement and (2) to discover if *Neotrypaea* sp. presence influence microbiome of the sediment itself. In this way, the bacterial community could act as both a cause and an effect of *Neotrypaea* sp. habitat utilization. The genera *Shewanella* sp. and *Pseudoalteromonas* sp. were more highly abundant in sediment associated with the new offshore population. *Shewanella* specifically was proportionally higher in *Neotrypaea* sp. positive sediment in both the offshore and estuarine habitat. Although offshore sediment and estuarine sediment have distinctive microbiomes, there are common bacteria among offshore and estuarine *Neotrypaea* sediment which was reflected in the ordination plot.

More sampling and processing across offshore and estuary sites would improve the robustness of this study. Due to machine issues and a barcoding error of 8 of the offshore samples resulting in 100k reads/sample instead of 937K reads/sample in the original runs. In addition, presence and absence designations were initially based on if box core brought up a *Neotrypaea*. However, camera imaging at the Neskowin site indicated that there were high abundances of *Neotrypaea* burrow openings at some stations that were initially labelled as absences (so most likely were false absences); these stations were re-designated as presences, reducing the overall number of "absence" samples for offshore.

Moving forward with the findings of this research, future directions should include settlement experiments with larvae at varying environmental conditions. Particularly, species of *Shewanella* sp. should be selected – such as *Shewanella colwelliana* which is known to promote oyster settlement in *Crassostrea gigas* and *C. virginica* (Weiner et al. 1989). Oyster larvae are

influenced both by the physical properties of the biofilm as well as chemical cues in the water column (Wassnig and Southgate 2012). As Neotrypaea *sp*. are a pest of oysters, *Shewenella*-induced settlement could have possible implications for oyster aquaculture management. Including experiments with temperature, oxygen, salinity, or pH fluctuations such as in the study by Whalan and Webster (2014) could also offer insight as to how environmental changes may be contributing to microbial shifts and to novel populations.

The overall microbial composition of sediment with burrowing shrimp has implications for sediment microbiome and nutrient cycling. As there are greater abundance of sulfate reducers with *Neotrypaea* sp. presence, the sulfur cycle would be heavily influenced. Sulfate reducers are known to compete with methanogens and overall could reduce methane emissions – as demonstrated with cable bacteria in rice paddies (Worner et al. 2019, Scholtz et al. 2020). Cable bacteria, as they can continuously couple sulfide oxidation with oxygen reduction, continuously replenish the sulfate available for other sulfate reducers to utilize and therefore compete with methanogens (Sandfield *et al*. 2020). Sequences of the bacteria of the gut and burrow of another burrowing shrimp, *Upogebia major*, were compared using this analysis in a previous study that focused on cable bacteria (Li et al. 2020). Although *Candidatus electrothrix* could be an indicator genus for *Neotrypaea sp*. and a genus that grows under bioturbation, cable bacteria have not before been associated with larval settlement.

There are still many more questions and experiments to pursue in order to fulfill the original objectives, but this research made progress. In addition to experiments involving settlement, the microbial composition of the sediments before and after *Neotrypaea* introduction should investigate burrowing effect on sulfate-reducers and cable bacteria over time.

5. References

Aller, J. Y., & Aller, R. C. (1986). Evidence for localized enhancement of biological associated with tube and burrow structures in deep-sea sediments at the HEEBLE site, western North Atlantic. Deep Sea Research Part A. Oceanographic Research Papers, 33(6), 755-790.

Bao, W. Y., Satuito, C. G., Yang, J. L., & Kitamura, H. (2007). Larval settlement and metamorphosis of the mussel *Mytilus galloprovincialis* in response to biofilms. Marine Biology, 150(4), 565-574.

Bertics, V. J., & Ziebis, W. (2009). Biodiversity of benthic microbial communities in bioturbated coastal sediments is controlled by geochemical microniches. The ISME journal, 3(11), 1269-1285.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods, 13(7), 581-583.

Cavalcanti, G. S., Alker, A. T., Delherbe, N., Malter, K. E., & Shikuma, N. J. (2020). The influence of bacteria on animal metamorphosis. Annual Review of Microbiology, 74(1).

DeWitt, T. H., D'Andrea, A. F., Brown, C. A., Griffen, B. D., & Eldridge, P. M. (2004). Impact of burrowing shrimp populations on nitrogen cycling and water quality in western North American temperate estuaries. In Proceedings of the symposium on "Ecology of large bioturbators in tidal flats and shallow sublittoral sediments-from individual behavior to their role as ecosystem engineers (pp. 107-118).

Dumbauld, B. R., & Bosley, K. M. (2018). Recruitment ecology of burrowing shrimps in US Pacific coast estuaries. Estuaries and Coasts, 41(7), 1848-1867.

Dumbauld, B. R., Chapman, J. W., Torchin, M. E., & Kuris, A. M. (2011). Is the collapse of mud shrimp (*Upogebia pugettensis*) populations along the Pacific coast of North America caused by outbreaks of a previously unknown bopyrid isopod parasite (*Orthione griffenis*)? Estuaries and Coasts, 34(2), 336-350.

Dobretsov, S., & Rittschof, D. (2020). Love at first taste: induction of larval settlement by marine microbes. International Journal of Molecular Sciences, 21(3), 731.

Espinel-Velasco, N., Tobias-Hünefeldt, S. P., Karelitz, S., Hoffmann, L. J., Morales, S. E., & Lamare, M. D. (2021). Reduced seawater pH alters marine biofilms with impacts for marine polychaete larval settlement. Marine Environmental Research, 167, 105291.

Feldman, K. L., Armstrong, D. A., Eggleston, D. B., & Dumbauld, B. R. (1997). Effects of substrate selection and post-settlement survival on recruitment success of the thalassinidean shrimp *Neotrypaea californiensis* to intertidal shell and mud habitats. Marine Ecology Progress Series, 150, 121-136.

Griffis, R. B., & Suchanek, T. H. (1991). A model of burrow architecture and trophic modes in thalassinidean shrimp (Decapoda: Thalassinidea). Marine Ecology Progress Series, 171-183.

Hadfield, M. G. (1984). Settlement requirements of molluscan larvae: new data on chemical and genetic roles. Aquaculture, 39(1-4), 283-298.

Hadfield, M. G. (2011). Biofilms and marine invertebrate larvae: what bacteria produce that larvae use to choose settlement sites. Annual review of marine science, 3, 453-470.

Huang, S., & Hadfield, M. G. (2003). Composition and density of bacterial biofilms determine larval settlement of the polychaete *Hydroides elegans*. Marine Ecology Progress Series, 260, 161-172.

Labare, M. P., & Weiner, R. M. (1990). Interactions between *Shewanella colwelliana*, oyster larvae, and hydrophobic organophosphate pesticides. Applied and environmental microbiology, 56(12), 3817-3821.

Leduc, D., & Pilditch, C. A. (2017). Estimating the effect of burrowing shrimp on deep-sea sediment community oxygen consumption. PeerJ, 5, e3309.

Li, C., Reimers, C. E., & Chapman, J. W. (2020). Microbiome analyses and presence of cable bacteria in the burrow sediment of *Upogebia pugettensis*. Marine Ecology Progress Series, 648, 79-94.

MacGinitie, G. E. (1934). The natural history of *Callianassa californiensis* Dana. American Midland Naturalist, 166-177.

Peng, L. H., Liang, X., Xu, J. K., Dobretsov, S., & Yang, J. L. (2020). Monospecific biofilms of Pseudoalteromonas promote larval settlement and metamorphosis of *Mytilus coruscus*. Scientific reports, 10(1), 1-12.

Rischer, M., Klassen, J. L., Wolf, T., Guo, H., Shelest, E., Clardy, J., & Beemelmanns, C. (2016). Draft Genome Sequence of Shewanella sp. Strain P1-14-1, a Bacterial Inducer of Settlement and Morphogenesis in Larvae of the Marine Hydroid *Hydractinia echinata*. Genome Announcements, 4(1), e00003-16.

Scholz, V. V., Meckenstock, R. U., Nielsen, L. P., & Risgaard-Petersen, N. (2020). Cable bacteria reduce methane emissions from rice-vegetated soils. Nature communications, 11(1), 1-5.

Strasser, K. M., & Felder, D. L. (1999). Settlement cues in an Atlantic coast population of the ghost shrimp *Callichirus major* (Crustacea: Decapoda: Thalassinidea). Marine Ecology Progress Series, 183, 217-225.

Wassnig, M., & Southgate, P. C. (2012). Effects of settlement cues on behaviour and substrate attachment of hatchery reared winged pearl oyster (*Pteria penguin*) larvae. Aquaculture, 344, 216-222.

Weiner, R. (1989). Effect of biofilms of the marine bacterium *Alteromonas colwelliana* (LST) on set of the oysters *Crassostrea gigas* (Thunberg, 1793) and C. virginica (Gmelin, 1791). J. Shellfish. Res., 8, 117-123.

Wörner, S., & Pester, M. (2019). Microbial succession of anaerobic chitin degradation in freshwater sediments. Applied and environmental microbiology, 85(18), e00963-19.

Whalan, S., & Webster, N. (2014). Sponge larval settlement cues: the role of microbial biofilms in a warming ocean. Scientific Reports, 4(1), 1-5.

Yang, J. L., Shen, P. J., Liang, X., Li, Y. F., Bao, W. Y., & Li, J. L. (2013). Larval settlement and metamorphosis of the mussel *Mytilus coruscus* in response to monospecific bacterial biofilms. Biofouling, 29(3), 247-259.

CHAPTER 3: SPECIES DISTRIBUTION MODELING OF OFFSHORE *NEOTRYPAEA* SP. USING BOOSTED REGRESSION TREES (BRT) AND MAXIMUM ENTROPY (MAXENT) MODELS

1. Introduction

A large *Neotrypaea* sp. population discovered offshore Newport, OR, in 2019 by Henkel *et al*. (2022) is hypothesized to have settled there in 2014-2015 based on cohort analysis. Although settlers previously had been found offshore, this was the greatest density of *Neotrypaea* (per box core grab), and the first multi-cohort population that was discovered. Both *Neotrypaea gigas* and *Neotrypaea californiensis* are described as estuarine species of burrowing shrimp (MacGinitie, 1934; Ricketts and Calvin, 1952). Therefore, this new population brings to question the true extent of their suitable habitat; in particular, whether they occur elsewhere offshore.

The occurrence of a population of *Neotrypaea* that settled in 2014 or 2015 coincides with the warm water anomaly of 2013-2015 (Figure 15). This marine heat wave, termed 'the warm blob', is associated with higher temperatures, lower chlorophyll-*a* in the winter/spring and higher in the summer/fall, lower nutrients, lower oxygen, and an increase in ocean acidification (Cavole et al. 2015). Northward range shifts of have been demonstrated for crustaceans offshore northern California, which are often at their range limit, included the spiny lobster (*Panularis interruptus*), chocolate porcelain crab (*Petrolisthes manimaculis*) and xantus swimming crab (*Portunus xantusii*); all of which have a pelagic larval stage (Sanford et al. 2019). This reflects the northward transport of larvae not only due to the higher sea surface temperatures being more habitable to warm temperatures species, but also due to changes in oceanographic currents (Schultz et al. 2011, Helmuth et al. 2016, Wernberg et al. 2013).

The newly discovered offshore shrimp population, and perhaps other offshore populations, could have upstream effects on the coastal pelagic and estuarine ecosystem. *Neotrypaea* sp. may restructure the offshore benthos, as they disrupt the sediments, resulting in shifts in the macrofaunal assemblages, and contribute to the caloric intake of predators such as crabs, fishes, and gray whales (Weitkamp et al. 1992; Darling et al. 1998; Dunham and Duffus 2001, Dumbauld et al. 2008).

Marine mammals demonstrate a preference for prey based on caloric content and abundance, and it was recently shown that gray whales make trade-offs between foraging and food quality. Gray whales have a higher probability of foraging associated with the lower-caloric but more abundant *Holmesimysis sculpti* over the higher-caloric *Neomysis rayii* (Hildebrand et al. 2022). *Neotrypaea* are orders of magnitude larger than mysids, so an abundant bed of them could provide sufficient caloric density for groups of whales to forage, as described by Weitkamp et al. (1992). If *Neotrypaea* populations are increasing offshore this may improve the survival of gray whale population and ameliorate health problems, as reduced body mass is associated with high levels of stress hormone which occurred in the gray whale population in 2018 (Lemos et al. 2021) and contributed to the mass mortality event in 2019 (NOAA 2021). An expanded range of *Neotrypaea* offshore would change the caloric content of the shelf benthos and could result in higher trophic level effects.

Species Distribution Models (SDMs) integrate both species occurrence data (either presence-background or presence-absence data) and environmental data to predict the probability of occurrence across a landscape (Elith and Leathwick 2009) and organisms' environmental preferences based on the degree to which environmental parameters contribute to the models.

Ecologically relevant predictors are critical to understanding why a species exists where they do; however, if the intent is to only predict where the species occurs and not habitat parameters driving the distribution the environmental variables need not to be as relevant (Elith and Leathwick 2009). This provides a strong tool to identify areas for conservation, environmental planning, and document rare species distributions (Guisan & Thuiller 2005). Species distribution models are most accurate in predicting the occurrence of species (53% of studies successfully predict occurrence), rather than abundance, population mean fitness, or genetic diversity (Lee-Yaw et al., 2022).

I hypothesize that the important environmental variables for *Neotrypaea* sp. distribution in general include sediment parameters, oxygen, salinity, and depth, as the shrimp predominantly occur in estuaries which are lower salinity and depth compared to offshore areas. *N. californiensis* prefer higher sand areas of the upper intertidal, and *N. gigas* generally occur in muddy sand of the upper middle intertidal (Griffis and Chavez 1988; and Jensen, 2014). Relevant predictors in terms of environmental changes during the marine heatwave of 2013-2015 include oxygen, pH, temperature, and current velocity.

MaxEnt uses machine learning to predict the probability of distribution, response curves of the species to environmental parameters, and the percent contribution of each parameter to the full model. This Maximum Entropy (MaxEnt) approach is a commonly used presence-only modeling technique that has contributed to designating environmental areas for conservation (Kaky et al. 2020, Warrant and Seifert 2011). Generally, MaxEnt is used to predict suitability of habitat rather than probability of occurrence. Concerns with presence-only modeling include the confounding effect of sampling effort, which may cause locations with fewer samples to appear

less densely populated rather than potentially under-sampled (Winship et al. 2020, Peel et al. 2019). As our data are based on box-core sampling, the likelihood of *Neotrypaea* sp. being at a sample location without being collected is low; however, our sampling effort is somewhat patchy. Further, as we have not found *Neotrypaea* sp. at sites that otherwise seem similar to where we have collected them, it will be important to include absence data.

Absence data can be incorporated with Boosted Regression Trees, which is a type of SDM that uses regression trees and boosting to fit simpler binary splits in data to a combined model (Breiman et al. 1984, Elith et al. 2008). The binary splitting of regression trees is beneficial for using many predictor variables, as is the case in this study, because the irrelevant predictors are avoided, and multiple types of predictor variables can be used (Elith et al. 2008). The relative contribution of each predictor variable to the full model can be included in the output, which is dependent on the amount of times the specific variable that is selected for splitting (Elith et al. 2008, Friedman 2001). In this context, the relative contribution of each predictor to the model prediction of *Neotrypaea* sp. presence can be used to characterize their fundamental niche – why do they exist where they do?

The primary objectives of this study were to (1) determine characteristics of offshore suitable habitat for the *Neotrypaea* sp., (2) determine where else there is a high probability of *Neotrypaea* sp. presence offshore the Pacific Northwest.

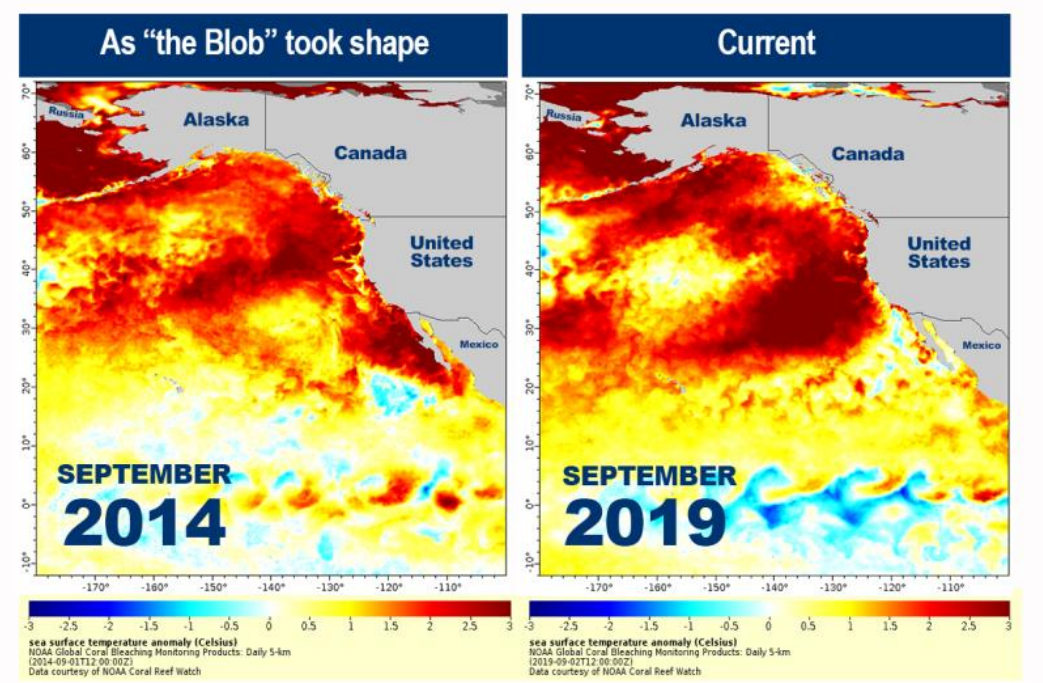


Figure 15. *Sea surface temperature anomaly from September 2014 (during the “warm blob”) and September 2019. NOAA Coral Reef Watch.*

Hypotheses:

(1) There is other highly suitable habitat for *Neotrypaea* sp. elsewhere on the shelf

(2) Environmental predictors related to the warm water anomaly (ocean temperature, chlorophyll, current velocity, oxygen) will have a higher relative influence in modeling shrimp occurrence than more constant variables associated with the seafloor conditions (benthic substrate type, rugosity, outcrop, slope).

2. Methods:

2.1 Input dataset

Neotrypaea presence and absence data were collated from an array of projects, including BOEM, EPA, CCLEAN, PMEC, OWET, SFPUC, NOAA, MorroBay, MCI (see Poti et al. 2021 for description of these data sources) in addition to more recent Henkel lab samples (PWS and PWN – see below) which were taken using a modified Grey-O'Hara 0.1 m^2 box corer and stored in 70% ethanol for sorting. Other *in situ* data collected included vertical water column profiles at many stations using a Sea-Bird Electronics CTD and sediment samples for grain size analysis. The PacWave South site (PWS) is located 13 km offshore of Newport Oregon at 70m depth and PacWave North (PWN) is located just over 3 km offshore at 40-50m depth. There were 39 samples taken at PWN between 2019 and 2020 and 98 at PWS sampled between 2019 and 2021. There were also 25 samples collected in October 2021 and April 2022 which were included in the final dataset for the final Hemery and BioOracle Models.

2.2 Model Runs

Species distribution models were created using environmental variables collected at each of the box core sampling stations (*in situ* data), coarse-grained Bio-Oracle environmental layers, environmental layers originally used for a BOEM OCS study (Poti et al. 2020; hereafter "Poti layers"), and environmental layers from modeling of sea star species on the west coast (Hemery et al. 2016; hereafter "Hemery layers") (Table 7). Environmental layers for the Maxent models included bio-ORACLE 2.1 benthic marine layers clipped to 39°N to 47°N and -128°W to -123°W in ArcGIS. Hemery and Bio-Oracle layers were used in presence-only MaxEnt models;

and *in situ*, Poti, Hemery, and Bio-Oracle layers were used in the BRT presence-absence models. All models were re-run with low-importance variables removed and compared to the original model run.

SDMs were conducted using the open-source software 'MaxEnt' (Phillips et al. 2006). Using MaxEnt, *Neotrypaea* sp. habitat suitability was modeled with occurrence data from 1999 to early 2022. This included a total of 350 presence points, 30 of which were held out for test data. Layers from Bio-Oracle included bottom temperature, salinity, dissolved oxygen, nitrate, phosphate, primary productivity, current velocity, silicate, dissolved molecular iron, chlorophyll, phytoplankton, and light at bottom

(Table 7). These layers have a coarse cell size, with a resolution of 5 arcmin or 9.2 km^2.

A MaxEnt Model was also run with the same occurrence dataset with Hemery layers; including temperature, salinity, eastward velocity, northward velocity, mean grain size, percent sand, percent mud, percent clay, percent

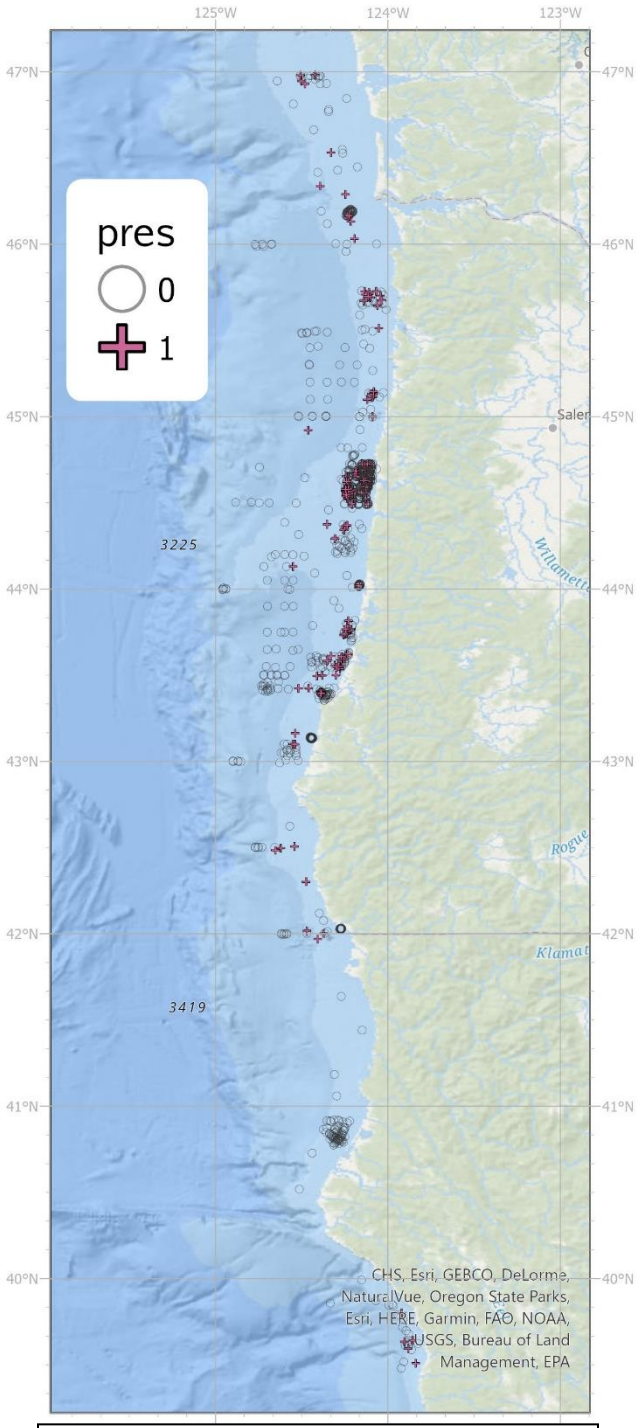


Figure 16. *Location of* Neotrypaea *data used as input into the species distribution models (1999 – April 2022).*

gravel, depth, slope, and rugosity at a cell size of 2 km^2 (Table 7). The correlation for annual salinity, annual temperature, and depth was > 0.90, but they were still included due to the ability of MaxEnt to withstand highly correlated variables (Hemery et al. 2016, Elith et al., 2011). MaxEnt models were run with 5 replicates of the cross-validate replicate run type and a convergence threshold of 0.00001.

Boosted Regression Tree Models were run with *in situ* data, Poti, Hemery, and Bio-Oracle layers. Poti environmental data were extracted from the predictor layers at the sample locations up to October 2019 by Matthew Poti, and environmental data layers were removed if correlated. The final environmental variables for the Poti layers included depth, aspect eastness, aspect northness, bottom current velocity eastness for spring summer, vertical bottom current velocity for fall, distance to shore, general curvature, hard soft, latitude, percent gravel, percent sand, plan curvature evans, profile curvature evans, slope, wave power annual max, wave power spring-summer (Table 7). General curvature, hard soft, distance to shore, plan curvature evans, profile curvature evans, wave power annual max, and wave power spring summer were removed for subsequent model runs. The Boosted Regression Tree models with Bio-Oracle and Hemery Layers were run with occurrence data from 1999 - 2022 (Figure 16). A BRT model with all Bio-Oracle environmental variables was run followed by a model run without light at the bottom, phosphate, and phytoplankton. In addition to a model with all environmental layers available for Hemery, a model was run without rocky outcrop, rugosity, and percent clay; and models with other combinations of environmental also layers were evaluated.

Tuning parameters were chosen by running models with multiple combinations of parameters and selecting for the model with the highest percent deviance explained. Options for learning

rate included 0.005, 0.001, 0.0005, 0.0001; options for bag fraction included 50 and 75, and options for tree complexity included 2, 3, 4, 5, and 10. Models were rerun selecting for a minimum of 1000 trees, as recommended by Elith et al. (2008). The highest performing models were selected based on AUC, PDE, and Moran's I value. AUC was used to compare across MaxEnt and BRT models, and PDE was used to compare within BRT models. Area under the curve of 'AUC' is based on true positive vs false positive rates. Ideally there would be only true positives which would indicated by an AUC of 1. However, often the model will not be completely accurate and there will be false positives in the testing data resulting in a lower AUC. Generally, an AUC of above .75 is considered good. The PDE, or percent deviance explained, indicate the ability of the model to accurately predict presence or absence based on either the training PDE (all of the data) or the cross validation PDE (mean over model iterations). Moran's I index is used to test for spatial autocorrelation, or how related a value in a cell is to the value next to it, and is a value between -1 and 1. An index value that is close to the maximum and minimum valuesindicatethat the spatial distribution of values (presence or absence of a cell) are not clustered (+1) or dispersed (-1) but random.

Table 7*. Environmental layers used from Bio-Oracle, Hemery et al. 2016, and Poti et al. 2020. ARMOR: Global Observed Ocean Physics Reprocessing; PISCES: Global Ocean Biogeochemistry Non-assimilative Hindcast; ORAP: Global Ocean Physics Reanalysis ECMWF; GlobColour: merging MERIS/MODIS/SeaWiFS.*

Bio-Oracle	Unit	Source	Year	Size
Temperature	°C	ARMOR	2000–2014	9.2 km
Salinity	PSS	ARMOR	2000–2014	9.2 km
Current velocity	m/s	ORAP	2000–2014	9.2 km
Nitrate	µmol/m^3	PISCES	2000–2014	9.2 km
Phosphate	µmol/m^3	PISCES	2000–2014	9.2 km
Silicate	µmol/m^3	PISCES	2000–2014	9.2 km
Dissolved molecular oxygen	µmol/m^3	PISCES	2000–2014	9.2 km
Dissolved Iron	µmol/m^3	PISCES	2000–2014	9.2 km
Chlorophyll	mg.m^3	PISCES	2000–2014	9.2 km
Phytoplankton	µmol/m^3	PISCES	2000–2014	9.2 km
Primary productivity	g.m^{-1}.day^{-1}	PISCES	2000–2014	9.2 km
Light at bottom	e/m^2 /year	GlobColour	2000–2014	9.2 km
Hemery				
Temperature	°C	Durski et al., 2015	2008 - 2011	2 km
Salinity	PSU	Durski et al., 2015	2008 - 2011	2 km
Eastward Velocity	m/s	Durski et al., 2015	2008 - 2011	2 km
Northward Velocity	m/s	Durski et al., 2015	2008 - 2011	2 km
Mean Grain Size	ϕ	usSEABED		2 km
Percent Sand	%	usSEABED		2 km
Percent Mud	%	usSEABED		2 km
Percent Clay	%	usSEABED		2 km
Percent Gravel	%	usSEABED		2 km
Depth	m	Multiple Public Sources (Hemery et al. 2016)		2 km
Slope	°	Multiple Public Sources (Hemery et al. 2016)		2 km
Rugosity		Multiple Public Sources (Hemery et al. 2016)		2 km
Outcrop		Goldfinger et al., 2014; SR/V Derek M. Baylis using the Reson 8101 (240 kHz) multibeam sonar	2010	2 km
Poti				
Aspect Eastness	radians/100m	Derived from Depth		25 x 25 m
Aspect Northness	radians/100m	Derived from Depth		25 x 25 m
Bottom Current Velocity Eastness SpringSummer	m/s	UCSC 31 year hindcast ocean circulation model	1980-2010	25 x 25 m
Bottom Current Velocity Vertical Fall	m/s	UCSC 31 year hindcast ocean circulation model	1980-2010	25 x 25 m
Depth	m	Multibeam Bathymetry,NOAA CRM Vol. 6:8, GEBCO_2014 grid version 20150318		25 x 25 m
Distance to Shore		Poti et al. gridded spatial datasets		25 x 25 m
General Curvature	radians/100m	Derived from Depth		25 x 25 m
Hard Soft	N/A	Sources described in Goldfinger et al. 2014; CSUMB Seafloor Mapping Lab, USGS State Mapping Project, and MLML Center for Habitat Studies		25 x 25 m
Latitude		Poti et al. gridded spatial datasets		25 x 25 m
Percent Gravel	%	USGS usSEABED, OSU, SCCWRP, BOEM		25 x 25 m
Percent Sand	%	USGS usSEABED, OSU, SCCWRP, BOEM		25 x 25 m
Plan Curvature Evans	radians/100m	Derived from Depth		25 x 25 m
Profile Curvature Evans	radians/100m	Derived from Depth		25 x 25 m
Slope	°	Derived from Depth		25 x 25 m
Wave Power Annual Max	watts/m	NOAA WAVEWATCH III 30 year hindcast Phase 2, US West Coast (1979–2009)	1979–2009	25 x 25 m
Wave Power SpringSummer	watts/m	NOAA WAVEWATCH III 30 year hindcast Phase 2, US West Coast (1979–2009)	1979–2009	25 x 25 m

2.3 Model Ground-truthing

In October 2021 and April, July, and September 2022, we sampled sites with predicted high and low shrimp suitability/probability, resulting in a total of 138 stations. These collections were then compared with model predictions to see how well the models performed. Sites were selected based on the output of all six models and the time allotted and included from North to South: OOI-WA, Grays Harbor, Grayland, Willipa, MCR, Nehalem, Neskowin, OtterCrest, NH Line, PacWave South, and Tillicum.

3. Results

BRT with in-situ

The BRT model run with the *in-situ* data had a training Percent Deviance Explained (PDE) of 53.09, an AUC of 0.833, and a Moran's I value of 0.0099 (Table 8). A similar model run without percent fines or mean grain size had a lower PDE (35.721) but the same AUC value. Variables with the highest relative influence included depth, followed by TOC, median grain size, mean grain size, percent fines, percent gravel, percent sand, and salinity (Figure 22, Table 9).

BRT with Poti

The BRT model run using the Poti environmental layers had a training Percent Deviance Explained of 50.157 and a CV PDE of 20.379. The AUC was 0.807 and the Moran's I value of 0.016 (Table 8). The variable with the highest relative influence was depth followed by distance to shore, latitude, percent sand, aspect eastness, aspect northness, percent gravel, and bottom current velocity spring summer, and plan curvature evans (Table 9). The partial dependence plots (Figure 20) showed the *Neotrypaea* occurred where there was low gravel and high sand, bottom currents moving westward at .04 m/s, westward aspect, and a plan curvature of > 0 (a convex seafloor).

The BRT model with Poti layers displayed aggregates of highly suitable habitat (over 0.70) at 124.057°W 45.947°N, north of Nehalem; 124.057°W 45.947°, south of Netarts Bay; 124.072°W 45.0944°N, south of Nestucca Bay; 124.092°W 44.965°N, north of Siletz Bay; 124.133°W 44.594°N, offshore Yaquina Bay; 124.304°W 44.498°N far offshore between Yaquina Bay and Alsea Bay. In southern Oregon, there was more highly suitable habitat 124.257°W 43.733°N,

north of Winchester Bay; 124.473°W 43.321°N, south of Coos Bay; and 124.414°W 42.076°N south of Gold Beach (Figure 17, Panel 3).

BRT with Hemery

The BRT Model with Hemery included environmental layers of temperature, salinity, velocity, depth, outcrop, rugosity, slop clay, gravel, median grain size, mud, and sand. In this model, the probability of presence ranged from 0.0057 to 0.9326. The Hemery BRT model had a training percent deviance explained of 26.84 and a CV percent deviance explained of 16.63. The AUC value for this model was 0.782 and a Moran's I-value of $1.554e^{-15}$ (Table 8). The most influential variable was rugosity at 1.00 followed by depth, temperature, slope, mean grain size, eastward velocity and northward velocity most influential variables, above .40). Percent clay, gravel, and sand all had a scaled relative influence below 0.30 (Table 9). The partial dependent plots showed that the *Neotrypaea* sp. occurred where rugosity was greater than 0.70, percent mud was between 40 and 60%, vertical current velocity was northward and greater than 0.04 m/s, velocity was westward and greater than 0.04 m/s, mean grain size was between 0 and 2 φ, depth was between 0 and 100 meters and peaked at 75 m, annual average temperature was between 9 and 10 ℃, and slope was either 0 or greater than 2° (Figure 18).

Areas with a high probability of presence in the BRT model run with the Hemery layers included 124.371°W 46.610°N, offshore Willapa Bay, WA; 124.071°W 45.747°N, north of Nehalem Bay, OR; 124.129°W 44.732°N north of Yaquina Bay, OR; and 124.231°W 44.5511°N south of Yaquina Bay. In southern Oregon, areas with high probability of occurrence was predicted at 124.314°W 44.365°N, south of Alsea Bay; 124.255°W 43.72°N, offshore Winchester Bay; 124.568°W 43.216°N, south of Coos Bay, and 124.608°W 42.408°N, offshore Gold Beach.

North California also had some highly suitable locations at 124.527°W 40.462°N, south of Humboldt Bay (Figure 17, Panel 2). Most of the high probability of occurrence areas were distributed sparsely across the extent of the map.

Maxent with Hemery

The MaxEnt Model with Hemery layers had an AUC value of 0.906 +/- 0.031. The permutation importance, or the drop in AUC resulting in the random rearrangement of the variable on the training data, demonstrated that the most important variable was temperature (23.5%) followed by percent clay (23.2%). Depth, salinity, and northward velocity were all above 5%. The marginal response curve for temperature displayed a peak around 8-9°C and a positive northward current velocity with increasing intensity. Low percent clay was also preferable, with a peak at 0 followed by a subsequent decline (Figure 24).

The MaxEnt Hemery Model displayed large areas of high probability of presence offshore Yaquina Bay, 124.072°W 45.166°N; offshore Nestucca Bay, 124.215°W 44.622°N; offshore Nehalem Bay , 124.083°W 45.673°N; north of Coos Bay 124.297°W 43.579°N' and offshore Gray's Harbor, 124.469°W 46.971°N. This model included a range of suitable habitats from 0.0057 to 0.9327. The area of highly suitable habitat offshore Yaquina Bay extended 17.5 miles, and offshore Nestucca bay highly suitable habitat extended to 10 miles (Figure 17, Panel 5).

BRT with Bio-Oracle

The BRT model with Bio-Oracle had a training percent deviance explained of 27.588 and a cross-validation percent deviance explained of 15.669; the AUC for the Bio-Oracle BRT Model was 0.696 and the Moran's I p-value was 0.4501 (Table 8). The most influential variables were

current velocity with a relative influence of 23.35% and iron with a relative influence of 18.27%. These were followed by salinity, nitrate, silicate, temperature, and dissolved oxygen over 5% (Table 9). The partial dependence plots for the top six most influential variables showed that the shrimp occurred where iron was above 0.0008 μmol/m^3, current velocity was above 0.10 m/s, nitrate is below 24 μmol/m^3, under 33 PSS, peaks at a temperature of 9.8°C, silicate greater than 30 μmol/m^3, almost all chlorophyll above 0 mg/m^3, and phytoplankton everywhere except for 1 μmol/m^3 (Figure 19).

The probability of occurrence for the ranged 0.04394 to 0.66071. The model predicted areas with a high probability of *Neotrypaea* sp. presence at, 44.544°N 124.211°W, south of the Yaquina Bay Estuary; 124.134°W 45.131°N, south of Nestucca Bay; 124.111°W 45.7005°N, offshore Nehalem Bay; 124.315°W 46.282°N, offshore Youngs Bay; 124.3741°W 46.535°N, south of Willapa Bay; and 124.5194°W 46.9698°N, offshore Grays Harbor (Figure 17, Panel 4).

Maxent with Bio-Oracle

The MaxEnt Model with Bio Oracle layers had an AUC value of 0.914 +/- 0.013. The permutation importance, or the importance of the variable to the model, showed that silicate and primary productivity were the highest drop in model performance (57.6% and 26.8% respectively, Table 9). Iron and current velocity were relatively low importance for model contribution – with a permutation importance of 1.2 and 0.7. The marginal response curves for this model indicated that current velocity, phosphate, and nitrate showed weak impact on the probability of presence in the model based on their values. Low iron and low silicate were preferable, with peaks at the bottom range of values for both (Figure 23).

This MaxEnt Model resulted in a range of probability of presence from 0.0439 to 0.6607. High areas of suitable habitat included offshore Grays Harbor down to Willapa Bay, south of Youngs Bay, offshore Nehalem Bay down to Neharts Bay, north of Yaquina Bay, north of Winchester Bay, south of Coos Bay, and north of Humboldt Bay. Much of the west coast had a suitability score > 0.75 up to 10 miles offshore and farther than that, the only area > 0.90 suitability offshore was 124.409°W 44.532N, at 22 miles offshore of Yaquina Bay.

Model Ground-Truthing

In comparing our 138 sampled stations to the predicted probabilities of occurrence in their respective model cells, the BRT model using the Poti environmental layers had the second highest true positive rate and highest true absent rate (Table 10). The Bio-Oracle BRT model had a high true absent rate (63%) but a low true positive rate (33%), and the Bio-Oracle MaxEnt Model had the highest true positive rate (75%). The Hemery BRT Model was had a low true positive rate (12%) but a fairly high true absent rate (63%) (Table 10). The Hemery MaxEnt Model was about 50% for both true positives and true absences. There were sites in Nehalem and Neskowin that had high numbers of *Neotrypaea* collected per box core grab. At Neskowin, there were 75.5 *Neotrypaea* collected with an overall average of 3.8 shrimp per box core, and up to 8.5 *Neotrypaea* per grab with *Neotrypaea*. At Nehalem there were 55.5 *Neotrypaea* collected, an average of 6.94 per box core, and up to 16 shrimp per box core with *Neotrypaea* (see Appendix Figures 8 and 9). Among the 138 sampled stations 34% (48 box cores) had *Neotrypaea*, although locations were selected for both presences and absences.

Table 8. Tuning parameters and test statistics of the BRT models with in-situ data and environmental layers from Poti, Bio-Oracle, and Hemery.

Tuning Parameter/Test Statistic	in Situ	Poti (25m)	Bio-Oracle (9.2km)	Hemery (2 km)
Tree Complexity	10	4	4	5
Learning Rate	0.005	0.01	0.005	5.00E-04
Bag Fraction	0.5	0.75	0.5	0.5
Number of Trees	3900	1250	550	8550
Training Percent Deviance Explained	53.09	50.157	27.588	26.84
Cross-Validation Mean Percent Deviance Explained	25.254	20.379	15.669	16.63
Cross-Validation SE Percent Deviance Explained	3.183	2.684	3.3511	2.0718
Internal AUC	**0.83**	**0.807**	0.696	0.782
Moran's I Value	0.0099	0.016	0.4501	1.554E-15

Table 9*. Relative influence of environmental variables in the boosted regression tree models for the in situ, Poti, Bio-Oracle, and Hemery layers and the permutation importance of variables in the MaxEnt models. Bolded Variables are the top 8 performing variables for which partial dependence plots (Figures 18-21) and marginal response curves (Figures 22 & 23) are provided.*

in Situ	
Environmental variable	**BRT**
Depth	32.08
Mean Grain Size	9.65
%Fines	11.68
TOC	7.25
Median Grain Size	7.25
%Sand	6
Temp	4.46
Salinity	4.43
%Gravel	3.97
Oxygen	3.16
Fluorescence	2.92
pH	2.62
TN	2.53
Turbidity	2.34
Chlorophyll	0.66

Poti	
Environmental variable	**BRT**
Depth	12.33
Distance to shore	14.18
Latitude	8.94
%Sand	8.93
Aspect Eastness	7.28
Aspect Northness	6.66
%Gravel	6.28
Bottom CurrentV East SpringSummer	5.65
Plan Curvature Evans	5.31
Slope	4.75
General Curvature	4.73
Profile Curvature Evans	4.29
Wave Power SpringSummer	3.97
Wave Power Annual Max	3.32
Bottom CurrentV Vertical Fall	3.31
Hard Soft	0

Bio-Oracle			
Environmental variable	**BRT**	**Environmental variable**	**MaxEnt**
Current Velocity	23.35	Silicate	57.6
Iron	18.27	Primary Productivity	26.8
Salinity	12.65	Temperature	4.1
Nitrate	11.71	Light at bottom	3
Silicate	8.46	Salinity	2.8
Temperature	7.33	Clorophyll	2.5
Dissolved Oxygen	5.46	Dissolved Oxygen	1.3
Clorophyll	3.55	Current Velocity	1.2
Phosphate	3.22	Iron	0.7
Primary Productivity	2.37	Nitrate	0
Phytoplankton	2.24	Phosphate	0
Light at bottom	1.35	Phytoplankton	

Hemery			
Environmental variable	**BRT**	**Environmental variable**	**MaxEnt**
Depth	15.02	Temperature	23.5
Temperature	10.72	%Clay	23.2
Rugosity	9.7	Depth	16.6
Slope	9.17	Salinity	12.1
Salinity	8.44	Northward V	5.5
Mean Grain Size	8.38	%Mud	5.2
Eastward V	8.33	Rugosity	4.2
%Mud	7.37	%Grav	3.1
Northward V	7.33	Rocky Outcrop	2.5
Rocky Outcrop	7.29	Eastward V	1.5
%Sand	5.06	%Sand	1.4
%Clay	2.79	Mean Grain Size	0.9
%Grav	0.4	Slope	0.3

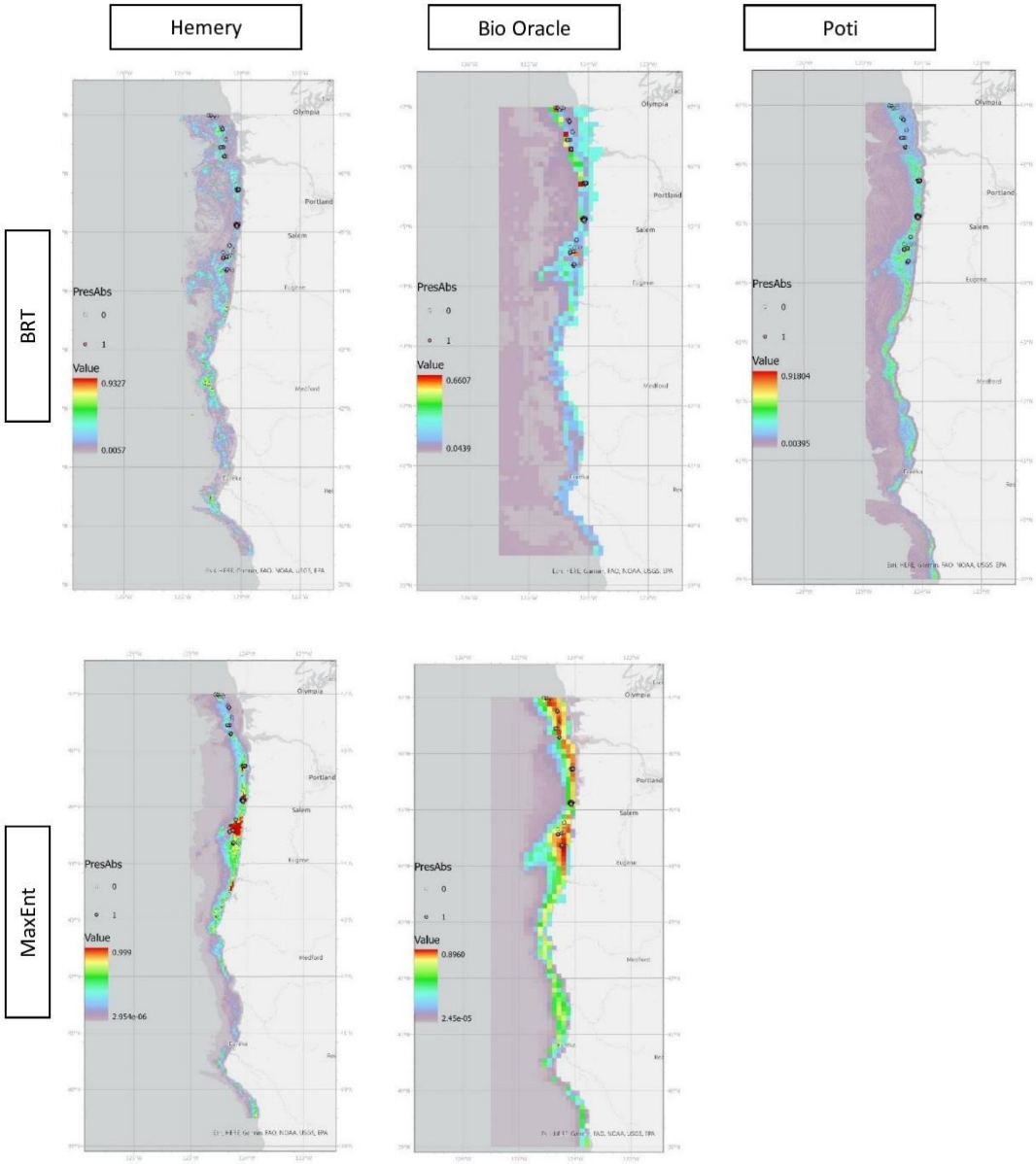


Figure 17. *Output of Species Distribution Models of* Neotrypaea *sp*. using *BRT (Top row) or MaxEnt (bottom row) with layers from Hemery, Bio-Oracle, and Poti. Presence Absence of Collections from the Community, Oxygen, and Productivity projects were overlayed on the Maps as well.*

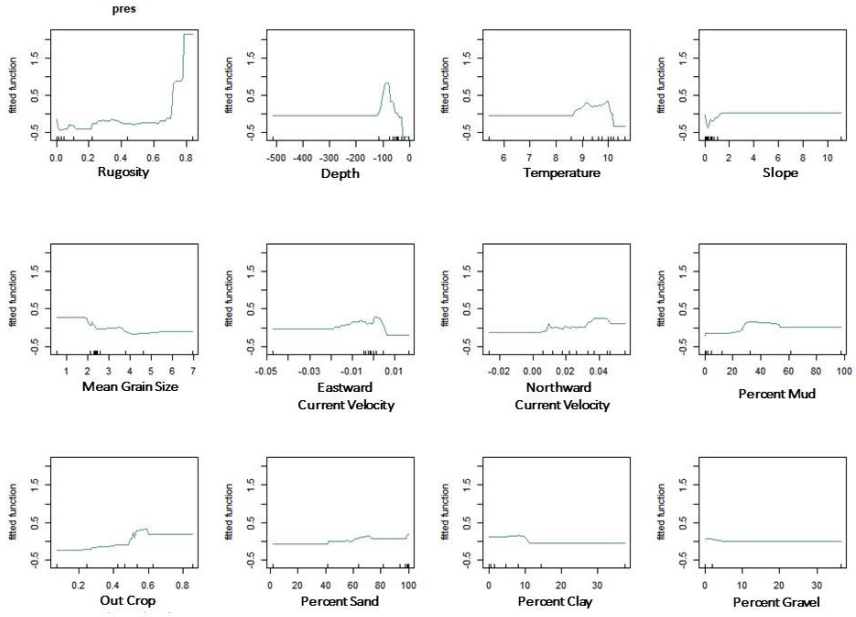


Figure 18. *Partial dependence with rug plots of the BRT Model with Hemery layers. Layers include Rugosity, Depth, Temperature, Slope (Top); Mean Grain Size, Eastward Current Velocity, Northward Current Velocity, Pereent Mud; Out Crop, Percent Sand, Percent Clay, Percent Gravel (Bottom)*

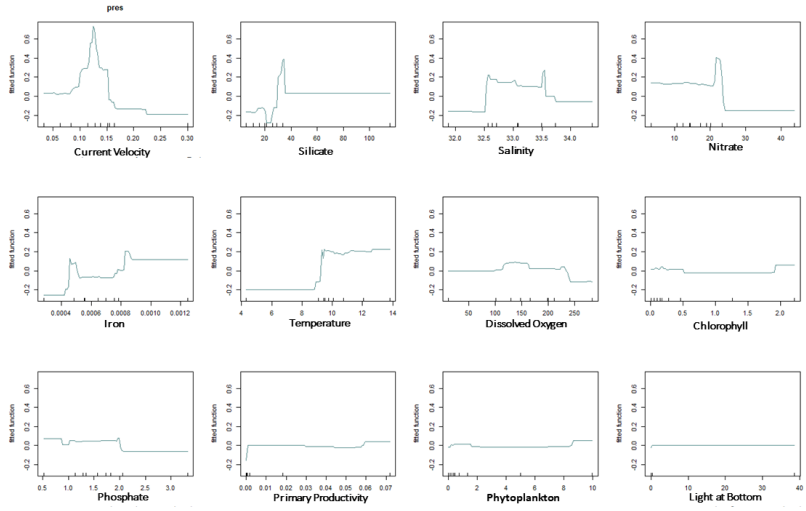


Figure 19. *Partial dependence and rug plots with BRT model with Bio-Oracle. Layers include Current Velocity, Silicate, Salinity, Nitrate (Top), Iron, Temperature, Dissolved Oxygen, Chlorophyll (Middle), Phosphate, Primary Productivity, Phytoplankton, Light at Bottom (Bottom).*

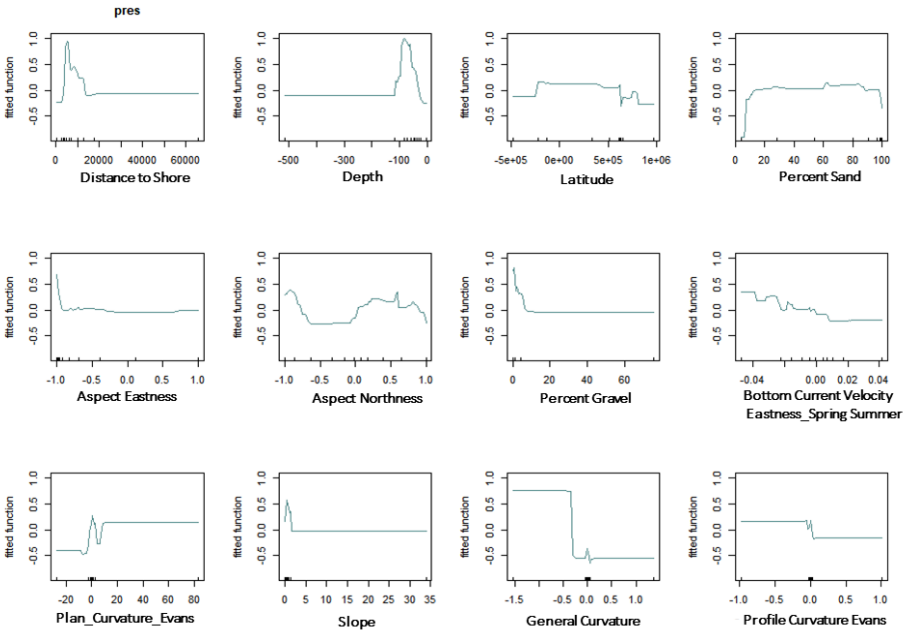


Figure 20. *Partial dependence and rug plots of BRT Model with Poti layers. Layers include Distance to Shore, Depth, Latitude, Percent Sand (Top), Aspect Eastness, Aspect Northness, Percent Gravel, Bottom Current Velocity for SpringSummer (Midddle), and Plan Curvature Evans, Slope, General Curvature, and Profile Curvature Evans (Bottom).*

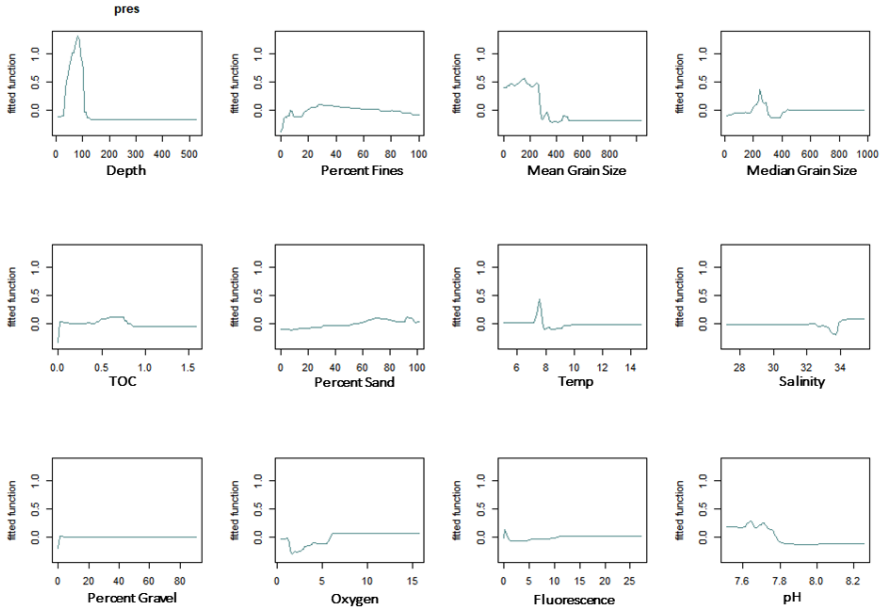


Figure 21. *Partial dependence plots with BRT Model with in situ data. Variables include Depth, Salinity, Mean Grain Size, Median Grain Size(Top); Total Organic Carbon, Percent Sand, Temp, Salinity (Middle); and Percent Gravel, Oxygen, Fluorescence, and pH (Bottom).*

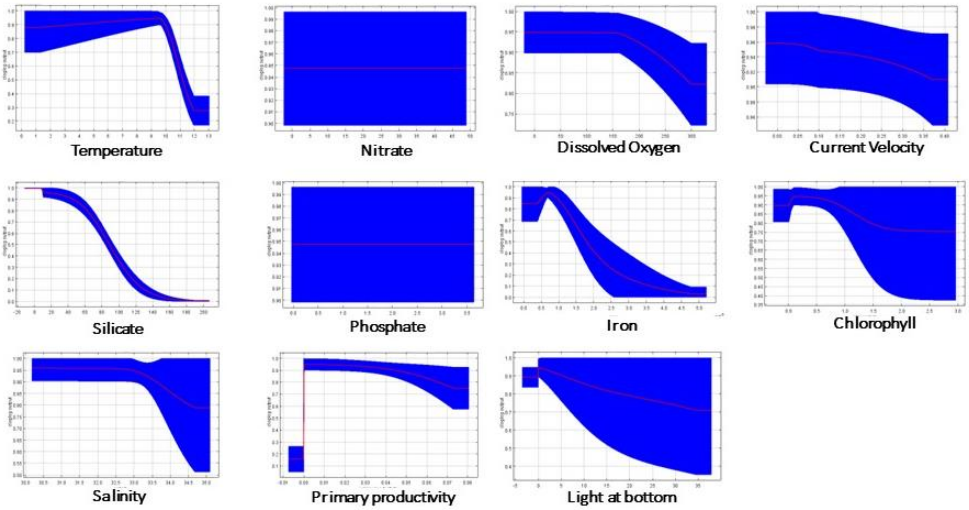


Figure 22. Marginal response curves for MaxEnt Bio-Oracle. Layers include Chlorophyll, CurrentVelocity, Dissolved Oxygen, Iron, Light at Bottom, Nitrate, Phosphate, Primary Productivity, Salinity, Silicate, and Temperature.

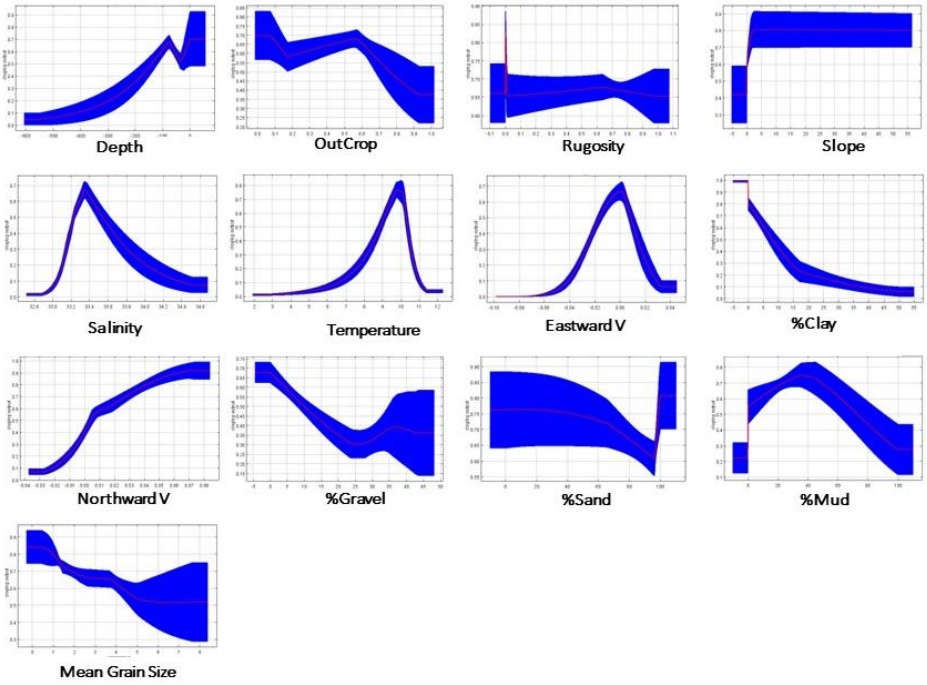


Figure 23. Marginal Response curves for MaxEnt Hemery. Layers Include Depth, Rocky OutCrop, Rugosity, Slope Annual Salinity, Annual Temperature, Eastward Current Velocity, Northward Current Velocity, Percent Clay, Percent Gravel, Mean Grain Size, Percent Mud, and Percent Sand.

Table 10. *True Presence, True Absence, False Absence, and False Positives from the ground-truthing sites and of the five models with spatial projections. Rates were calculated based on the total of absence or presences (TP/TP+FP). Thresholds were based on average value for the sampled stations and the histrogram of the distribution of total values for in ArcGIS Pro.*

	Threshold	True Pres	True Abs	False Abs	False Pos	TP Rate	TA Rate	FA Rate	FP Rate
BioOracle_BRT	a < .40 < p	14	60	29	35	0.33	0.63	0.67	0.37
BioOracle_Maxent	a < .701 < p	36	16	12	74	0.75	0.57	0.25	0.43
Hemery_BRT	a < .40 < p	3	71	22	42	0.12	0.63	0.88	0.37
Hemery_MaxEnt	a < .68 < p	22	50	23	46	0.49	0.52	0.51	0.48
Poti_BRT	a < .40 < p	17	76	17	28	0.50	0.73	0.50	0.27

4. Discussion

The two types of species distribution models used in this study (Maxent and BRT) offered insight into suitable habitat for *Neotrypaea* outside of their expected distribution in estuaries and indicated environmental conditions that could have occurred to initiate a bed of *Neotrypaea* offshore. The top two models that overall performed the best based on PDE were the boosted regression tree models using the *in situ* data and the Poti layers. PDE demonstrates the percent of the variance of the response variable (a presence or absence) at a given point. Precise, or highly resolute environment input, would then subsequently result in a higher PDE. Therefore, the resolution of the Poti layers (25m^2 grid cell size) was at a finer scale than the other layers, and so this was to be expected. The *in situ* model would also perform well as the environmental data precisely describes the locations collected.

Across the three boosted regression tree models that had depth as a potential predictor, this was one of the most influential of variables. *Neotrypaea* were more likely to occur in shallow waters, with a peak for an offshore population around 60-90m. In the best performing models, (*in situ* and Poti layers) sediment parameters were highly influential. Based on the partial dependence plots from both models, substrate with at least 60% sand and 40% mud, and areas with under 10% gravel or clay are most suitable for *Neotrypaea*. *Neotrypaea gigas* tend to be in areas of the mudflat with higher mud content than *Neotrypaea californiensis* (Griffis et al., unpubl. data; Mendez & Nguyen, 1987, Griffis and Chavez, 1988). Therefore, the mud and sand content are important where the *Neotrypaea* are found offshore and which species are found at specific locations. The use of all *Neotrypaea* (*N. californiensis*, *N. gigas*, and many unidentified) occurrences in the original model could be influencing this outcome. The variation of presences found in sand or mud could be from the collated data with all *Neotrypaea* offshore samples.

In the model using the *in-situ* data, total organic carbon was the third most influential variable. A study by Ferraro and Cole investigated similar environmental variables in habitat types in Willapa Bay and Grays Harbor, WA including *Neotrypaea* habitat (2011). These habitat types were organized based on a dominant species, including *Neotrypaea californiensis*. Out of all the estuarine habitats in those two estuaries, the sediment of *Neotrypaea californiensis*-dominated habitat showed the lowest percent TOC at 0.12% compared *Zostera japonica* habitat, with the highest percent TOC of 2.6% for comparison(Ferraro and Cole, 2011). For the *in situ* data, the marginal response curve for TOC showed a peak at about 0.10 and a decrease until 0.50. This may be because of detritus consumption and oxygen consumption by *Neotrypaea sp.*, taking in the organic carbon in the sediment and releasing CO^2. TOC is often used as a benthic

sediment indicator and is correlated with percent silt and benthic organisms (Longbottom, 1970; Santhanam, 2009; Henkel and Politano 2017).

I hypothesized that major drivers for *Neotrypaea* distribution would include oxygen and temperature. However, the relative importance of environmental variables associated with water conditions were different among model types. In the BRT and MaxEnt models where seawater conditions were included (*in situ*, Hemery, Bio-Oracle), temperature and/or salinity were at least in the top half of influential variables for the BRT models and. The *in situ* model response curve for salinity showed preferable values above 34 PSU while the Bio-Oracle Model showed a high prevalence of shrimp under 33.5 PSU. This contrast could be related to the difference between collected *in situ* data versus interpolated annual data. Locations close to the estuary and riverine input would have lower salinity values than locations further offshore – in Yaquina Bay, this ranges from 10 – 34 PSU. During the warm water anomaly in the northern California current, the upper 50-80m of water displayed temperature and salinity values close to freshwater (10-11 C and 31 PSU) (Peterson et al. 2015).

The annual average temperature for the model stretches over the range collected annually and reflects changes in temperature in the winter and summer months. On the NH line, the average winter temperature is 10.586 and the average summer temperature is 7.983 (depth of 50m, NOAA 2022). In August 2014 and August – October 2015, the anomalous SST on the Oregon Coast ranged from 1 – 2℃ above average (Cavole et al. 2016), Most of the *in situ* data were collected during the summertime, and the *in situ* temperature marginal response curve showed a large peak just before the summer average and then a steady increase to a plateau of the winter average and warmer. The Bio-Oracle BRT model mimicked the *in situ* response curve,

with a peak just before 10℃ and then a plateau after 12℃. The preferable temperature range for the Hemery BRT model was 9-10 degrees℃. Overall, temperatures collected at PWS ranged from 7.3 to 9.3℃ between 2013 and 2019 and in contrast the temperature of the *Neotrypaea* home estuary in Yaquina Bay ranges from 10 – 16℃ (Henkel et al. 2022). This may suggest that with the warm water anomaly and general increasing temperature, the offshore water temperatures can increase to that of the estuary. However, SST in the interpolated annual data would not be season-specific and CTD data is clumped during the times of collection so overall conclusions about temperature are limited.

The in situ data also showed a response curve of a higher probability of occurrence when oxygen is under 2 μmol/m^3 (hypoxic) and the peak for the bio-oracle layers ranged from 1.696 to 2.827 μmol/m^3. It has been shown before that *Neotrypaea* are tolerant to low oxygen (Thompson and Pritchard 1969), but also the i*n situ* results may be influenced by higher sampling during summer months. Primary productivity and chlorophyll were not highly influential to the model in the *in situ* BRT model or the bio-oracle BRT model.

A possible ecological explanation for the offshore population of *Neotrypaea* sp. and their existence offshore Nehalem and Nestucca are the riverine inputs. The larvae themselves are released into the ocean from the estuaries and so strong forcing may influence distance offshore or duration in pelagic waters. Iron, macronutrients, organic carbon, and larvae are fed to the coastal shelf offshore Oregon by riverine inputs. Downwelling prevents the dispersal of nutrient-rich water from rivers offshore, but wind relaxation or upwelling contributes to river plume extension (Wetz et al. 2006). Wetz suggested that if organic matter, iron, and macronutrients were able to extend further offshore during favorable winter conditions, this would in turn

promote productivity during the summer months and possibly alter the microbial food web. Winds during El Nino generally strengthens downwelling and limit offshore plumes.

Both the BRT and MaxEnt models predict the probability of presence (on a scale of 0 to 1), although the sampled stations did not reflect the probability of the presence of the models. However, the models did give insight into desirable offshore habitat for *Neotrypaea* and the potential for recurring populations. Over 70 *Neotrypaea* sp. were found offshore Neskowin, and over 50 *Neotrypaea* sp. offshore Nehalem in central Oregon. Prior to the 2019 discovery of offshore *Neotrypaea*, *Neotrypaea* were found in 15% of grabs in the collated dataset – most of which were 1 or 2 shrimp (Henkel et al. 2022). Out of the 138 samples collected for this project (some of which were targeted absence sites using the species distribution models), 36.1% had *Neotrypaea* present.

Future work could involve the creation of environmental layers from *in situ* data, and models with warming scenarios. Although we did not find large, recurring populations of *Neotrypaea* sp.in many areas of predicted high suitability, there is habitat available and sites where beds of burrowing shrimp occur. This research prompts more questions regarding the association of *Neotrypaea* sp. with mudflats and estuaries and brings to question their role of ecosystem engineers in the offshore benthic environment.

5. References

Bosley, K. M., Copeman, L. A., Dumbauld, B. R., & Bosley, K. L. (2017). Identification of burrowing shrimp food sources along an estuarine gradient using fatty acid analysis and stable isotope ratios. Estuaries and Coasts, 40(4), 1113-1130.

Brown, CA. Temperature, salinity and depth data from Yaquina Bay, Oregon from May 2012 through July 2012. U.S. Environmental Protection Agency, Washington, DC, 2019. Retrieved from https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=CPHEA&dirEntryId=347535

Brodie, S. J., Thorson, J. T., Carroll, G., Hazen, E. L., Bograd, S., Haltuch, M. A., ... & Selden, R. L. (2020). Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography, 43(1), 11-24.

Cavole, L. M., Demko, A. M., Diner, R. E., Giddings, A., Koester, I., Pagniello, C. M., ... & Franks, P. J. (2016). Biological impacts of the 2013–2015 warm-water anomaly in the Northeast Pacific: winners, losers, and the future. Oceanography, 29(2), 273-285.

Darling, J. D., Keogh, K. E., & Steeves, T. E. (1998). Gray whale (*Eschrichtius robustus*) habitat utilization and prey species off Vancouver Island, BC. Marine Mammal Science, 14(4), 692-720.

Dumbauld, B. R., Holden, D. L., & Langness, O. P. (2008). Do sturgeon limit burrowing shrimp populations in Pacific Northwest estuaries? Environmental Biology of Fishes, 83(3), 283-296.

Dunham, J. S., & Duffus, D. A. (2001). Foraging patterns of gray whales in central Clayoquot Sound, British Columbia, Canada. Marine Ecology Progress Series, 223, 299-310.

DeWitt, T. H., D'Andrea, A. F., Brown, C. A., Griffen, B. D., & Eldridge, P. M. (2004). Impact of burrowing shrimp populations on nitrogen cycling and water quality in western North American temperate estuaries. In Proceedings of the symposium on "Ecology of large bioturbators in tidal flats and shallow sublittoral sediments-from individual behavior to their role as ecosystem engineers (pp. 107-118).

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics, 40(1), 677-697.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. Diversity and distributions, 17(1), 43-57.

Escalante, T., Rodríguez-Tapia, G., Linaje, M., Morrone, J. J., & Noguera-Urbano, E. (2014). Mammal species richness and biogeographic structure at the southern boundaries of the Nearctic region. Mammalia, 78(2), 159-169.

Ferraro, S. P., & Cole, F. A. (2011). Ecological periodic tables for benthic macrofaunal usage of estuarine habitats in the US Pacific Northwest. Estuarine, Coastal and Shelf Science, 94(1), 36-47.

Fitzgerald, M., Coulson, R., Lawing, A. M., Matsuzawa, T., & Koops, K. (2018). Modeling habitat suitability for chimpanzees (*Pan troglodytes verus*) in the Greater Nimba Landscape, Guinea, West Africa. Primates, 59(4), 361-375.

Fourcade, Y., Engler, J. O., Rödder, D., & Secondi, J. (2014). Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. PloS one, 9(5), e97122.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. Ecology letters, 8(9), 993-1009.

Griffis, R. B., & Chavez, F. L. (1988). Effects of sediment type on burrows of *Callianassa californiensis* Dana and *C. gigas* Dana. Journal of Experimental Marine Biology and Ecology, 117(3), 239-253.

Helmuth, B., Mieszkowska, N., Moore, P., & Hawkins, S. J. (2006). Living on the edge of two changing worlds: forecasting the responses of rocky intertidal ecosystems to climate change. Annual review of ecology, evolution, and systematics, 373-404.

Hemery, L. G., Marion, S. R., Romsos, C. G., Kurapov, A. L., & Henkel, S. K. (2016). Ecological niche and species distribution modelling of sea stars along the Pacific Northwest continental shelf. Diversity and Distributions, 22(12), 1314-1327.

Henkel, SK and Politano, KK. 2017. Small proportions of silt linked to distinct and predictable differences in marine macrofaunal assemblages on the continental shelf. Continental Shelf Research 144: 38 – 49. https://doi.org/10.1016/j.csr.2017.06.016

Henkel SK and Gilbane LG. 2020. Using benthic macrofaunal assemblages for a bottom-up approach to defining habitat types in the NE Pacific sedimentary shelf and slope. Estuarine, Coastal, and Shelf Science. 246:107056. c

Henkel, SK, Revelas, EC, Wodzicki, S, Chapman, J. 2022. Discovery of a large offshore population of the northeast Pacific burrowing shrimp *Neotrypaea* sp. (Decapoda: Axiidea). Estuarine Coastal and Shelf Science. 274:107936. https://doi.org/10.1016/j.ecss.2022.107936

Hickey, B. M., & Banas, N. S. (2003). Oceanography of the US Pacific Northwest coastal ocean and estuaries with application to coastal ecology. Estuaries, 26(4), 1010-1031.

Hildebrand, L., Sullivan, F. A., Orben, R. A., Derville, S., & Torres, L. G. (2022). Trade-offs in prey quantity and quality in gray whale foraging. Marine Ecology Progress Series, 695, 189-201.

Kaky, E., Nolan, V., Alatawi, A., & Gilbert, F. (2020). A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: A case study with Egyptian medicinal plants. Ecological Informatics, 60, 101150.

Lee-Yaw, A., J., L. McCune, J., Pironon, S., & N. Sheth, S. (2022). Species distribution models rarely predict the biology of real populations. Ecography, 2022(6), e05877.

Lemos, L. S., Olsen, A., Smith, A., Burnett, J. D., Chandler, T. E., Larson, S., ... & Torres, L. G. (2022). Stressed and slim or relaxed and chubby? A simultaneous assessment of gray whale body condition and hormone variability. Marine Mammal Science, 38(2), 801-811.

Longbottom, M. R. (1970). The distribution of *Arenicola marina* (L.) with particular reference to the effects of particle size and organic matter of the sediments. Journal of Experimental Marine Biology and Ecology, 5(2), 138-157.

MacGinitie, G. E. (1934). The natural history of *Callianassa californiensis* Dana. American Midland Naturalist,

166-177.

Phillips, S. J. (2005). A brief tutorial on Maxent. AT&T Research, 190(4), 231-259.

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. Ecological modelling, 190(3-4), 231-259.

Poti, M., Henkel, S. K., Bizzarro, J. J., Hourigan, T. F., Clarke, M. E., Whitmire, C. E., ... & Jeffrey, C. F. G. (2020). Cross-Shelf Habitat Suitability Modeling: Characterizing Potential Distributions of Deep-Sea Corals, Sponges, and Macrofauna Offshore of the US West Coast. Camarillo (CA): US Department of the Interior, Bureau of Ocean Energy Management. OCS Study. BOEM, 21, 267.

NOAA Fisheries. 2021. 2019-2022 Gray Whale Unusual Mortality Event along the West Coast and Alaska. Retrieved from https://www.fisheries.noaa.gov/national/marine-life-distress/2019-2022-gray-whale-unusual-mortality-event-along-west-coast-and.

NOAA Fisheries. 2022. Newport Line Oceanographic Temperature Data. Retrieved from https://www.fisheries.noaa.gov/data-tools/newport-line-oceanographic-temperature-data.

Nguyen, D., & Leung, B. (2022). How well do species distribution models predict occurrences in exotic ranges? Global Ecology and Biogeography, 31(6), 1051-1065.

Ricketts, E. F., and J. Calvin. 1952. Between Pacific tides: an account of the habits and habitats of some five hundred of the common, conspicuous seashore invertebrates of the Pacific Coast between Sitka, Alaska, and Northern Mexico. Stanford: Stanford University Press, Stanford.

Saldías, G. S., Strub, P. T., & Shearman, R. K. (2020). Spatio-temporal variability and ENSO modulation of turbid freshwater plumes along the Oregon coast. Estuarine, Coastal and Shelf Science, 243, 106880.

Sanford, E., Sones, J. L., García-Reyes, M., Goddard, J. H., & Largier, J. L. (2019). Widespread shifts in the coastal biota of northern California during the 2014–2016 marine heatwaves. Scientific reports, 9(1), 1-14.

Santhanam, H. (2009). Influence of the sediment composition and total organic carbon on distribution of benthic organisms at Pulicat Lagoon, South East coast of India. Journal of Ecotoxicology and Environmental Monitoring, 19, 541-548.

Schultz, S. T., Goddard, J. H., Gosliner, T. M., Mason, D. E., Pence, W. E., McDonald, G. R., ... & Pearse, J. S. (2011). Climate-index response profiling indicates larval transport is driving population fluctuations in nudibranch gastropods from the northeast Pacific Ocean. Limnology and oceanography, 56(2), 749-763.

Thompson, R. K., & Pritchard, A. W. (1969). Respiratory adaptations of two burrowing crustaceans, *Callianassa californiensis* and *Upogebia pugettensis* (Decapoda, Thalassinidea). The Biological Bulletin, 136(2), 274-287.

Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W., Edgar, G. J., Mouillot, D., ... & Pellissier, L. (2022). A quantitative review of abundance-based species distribution models. Ecography, 2022(1).

Warren, D. L., & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological applications, 21(2), 335-342.

Weitkamp, L. A., Wissmar, R. C., Simenstad, C. A., Fresh, K. L., & Odell, J. G. (1992). Gray whale foraging on ghost shrimp (*Callianassa californiensis*) in littoral sand flats of Puget Sound, USA. Canadian Journal of

Zoology, 70(11), 2275-2280.

Wetz, M. S., Hales, B., Chase, Z., Wheeler, P. A., & Whitney, M. M. (2006). Riverine input of macronutrients, iron, and organic matter to the coastal ocean off Oregon, USA, during the winter. *Limnology and Oceanography*, *51*(5), 2221-2231.

Wernberg, T., Bennett, S., Babcock, R. C., De Bettignies, T., Cure, K., Depczynski, M., ... & Wilson, S. (2016). Climate-driven regime shift of a temperate marine ecosystem. Science, 353(6295), 169-172.

CONCLUSION

Morphological, microbial, and modeling research offered insight into the ecology of the offshore population of *Neotrypaea* sp. These three topics coalesce around fundamental questions brought up by the new, recurring population of *Neotrypaea* sp. Who are the offshore *Neotrypaea*? Why are they there? Where else could they be? What will they do to their environment? The insights of this research shed light on many aspects of the offshore population; such as identification, ideal habitat conditions, contributing environmental conditions to distribution and settlement, and their influence on sediment.

Morphological analysis showed that the offshore population was not identical to the estuarine species of *Neotrypaea gigas* as initially predicted. There were some novel characteristics distinguishing the offshore population from the known estuarine species, particularly the size of the corneas. The linear discriminant analysis of multiple measurements indicated that the three populations were distinct. This could indicate a hybrid species offshore, or a population of *N. gigas* with some morphological variation. The latter is most likely, but more genetic analysis should be conducted.

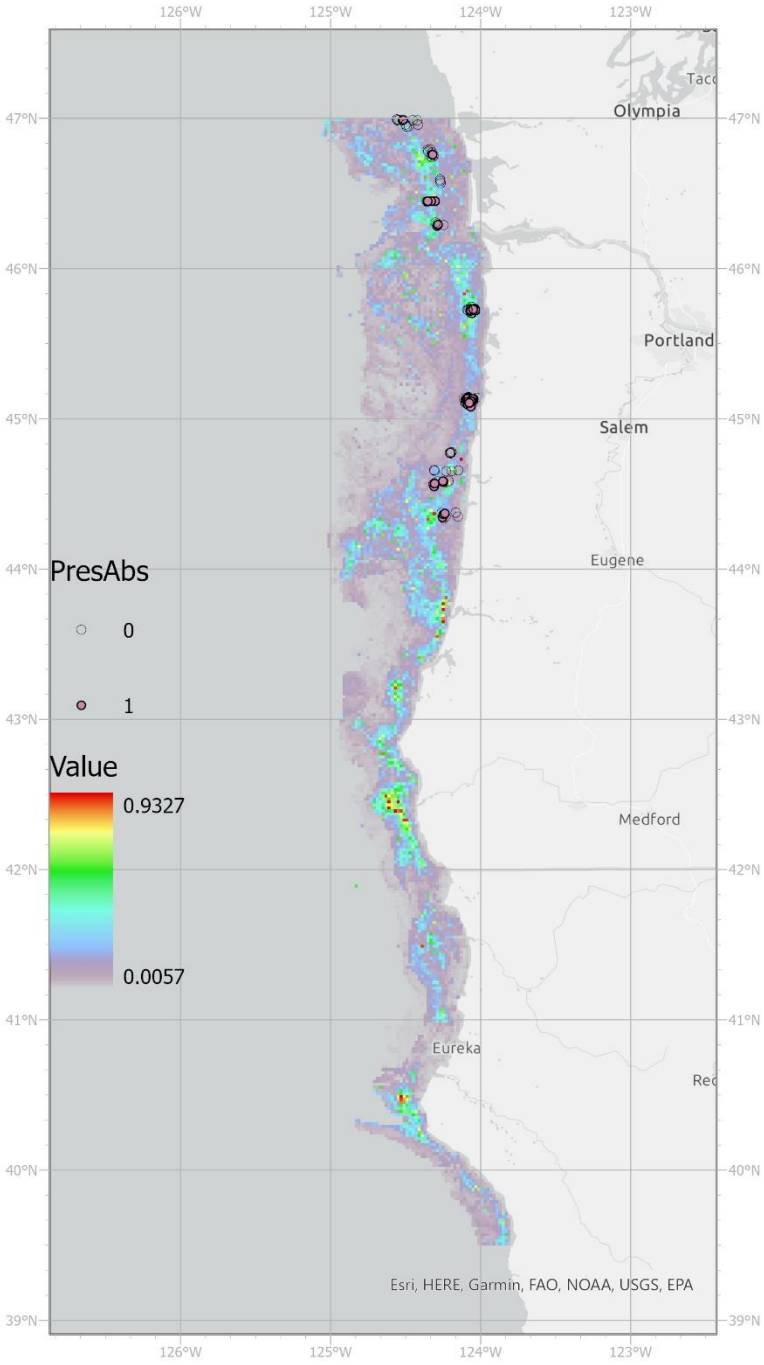
The microbial analysis demonstrated that there were greater abundances of *Shewanella* sp. in both offshore and estuarine species with *Neotrypaea*. This may have been settlement cue for *Neotrypaea* zoea or decapodids offshore prior to returning to the estuary. In turn, the presence of Neotrypaea resulted in changes to the microbial structure of the sediment they inhabited in both the estuary and offshore environments. Sulfate-reducers and sulfide oxidizers (e.g. cable bacteria) are descriptive for sediments containing *Neotrypaea* sp. both offshore and in the estuary.

Species Distribution Modeling indicated that the continental shelf offshore Washington, Oregon, and Northern California had areas of high probability of presence, and ground-truthing showed Neskowin and Nehalem, OR to host population of *Neotrypaea*. The results of the models followed the common axiom that no model is correct but some are useful. The BRT Poti model had the highest true absent rate and second highest true positive rate. The relative influence of each environmental variable, and response curves of the model to them, gave insight and description to their ideal habitat conditions. Sediment and grain size characteristics were important across models – particularly percent sand. This could be a reflection of the importance of non-organic settlement cues for zoea and decapodid burrowing as described by Strasser and Feldman (1999). Current velocity strength and direction (eastward) in the BRT Poti model reflects the potential importance of a changing environmental variable (current) in settlement of the offshore population.

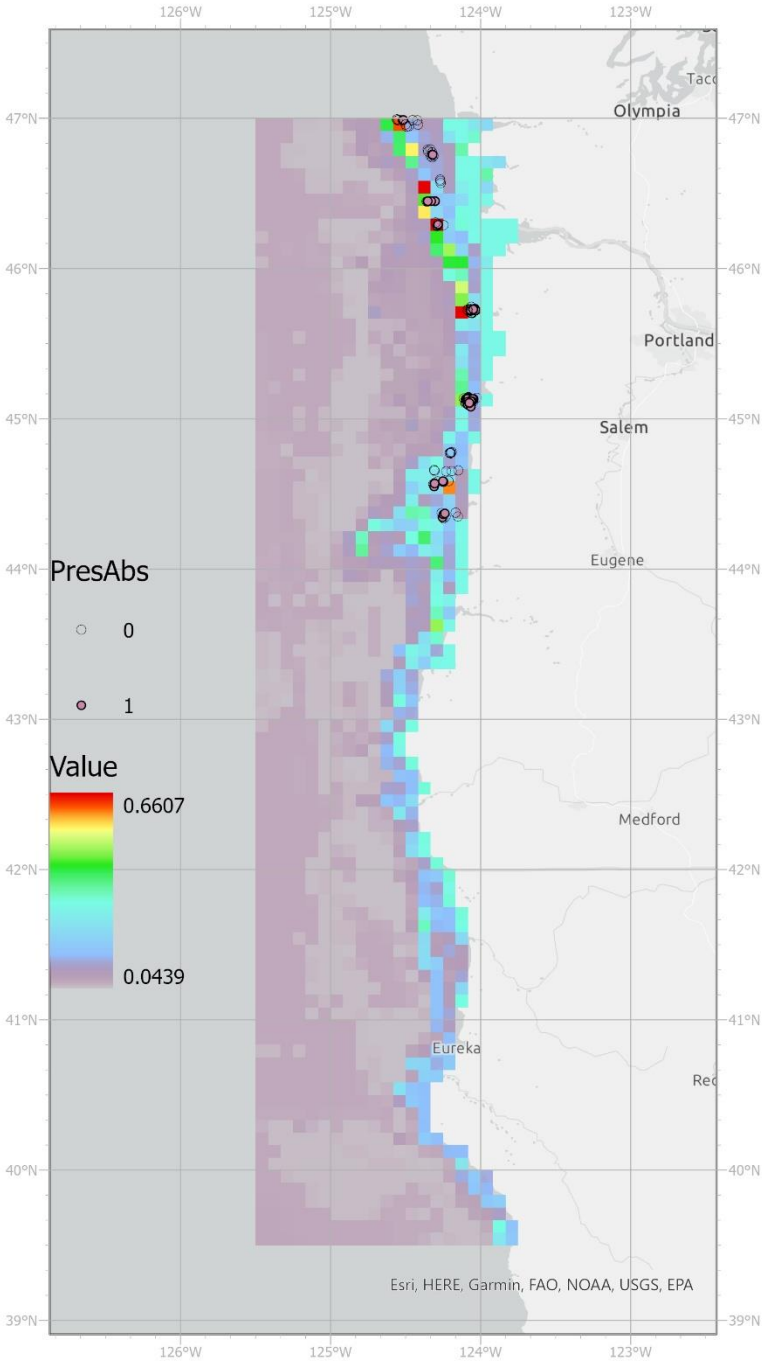
Why did the offshore population occur? Sediment characteristics have not changed at the time of settlement for the new offshore population. This research indicates that there may have been strong currents moving eastward that deflected the novel population, environmental changes that could have triggered growth of settlement-cueing bacteria in the benthos, and/or morphological differences that created competitive disadvantage in returning to the estuary or in successfully establish themselves offshore. Another consideration not included in the scope of this work include that warming temperatures of the warm water anomoly may have triggered early development and settlement of the larvae. Further research should target settlement of larvae to microbial biofilms and sediment types under varying environmental conditions.

Appendix: Supplemental Figures and Tables

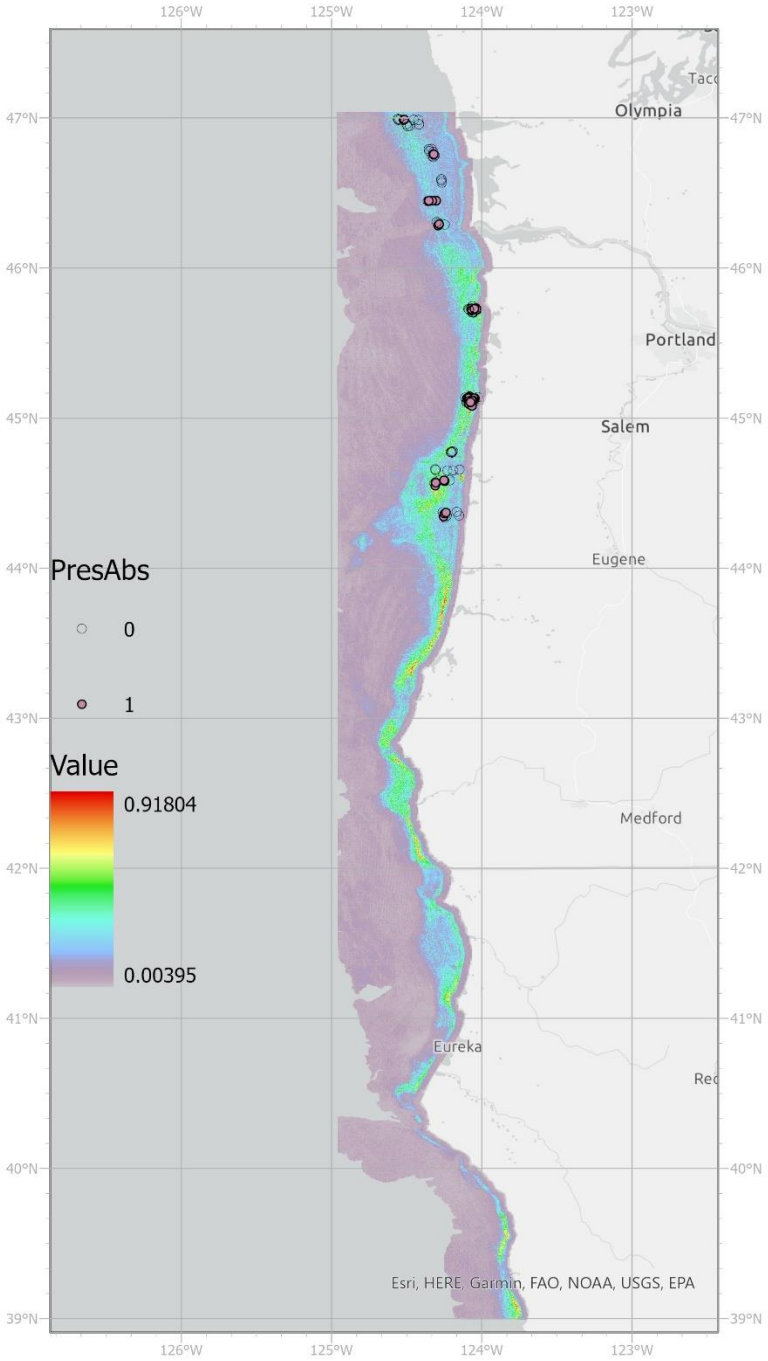
A1. Complete appendage images of the major claws of the offshore *Neotrypaea* (first), the estuarine *N. gigas* (middle) and the estuarine *N. californiensis* (last). This prompted the initial observation that the offshore *Neotrypaea* sp. were most likely estuary *N. gigas*.



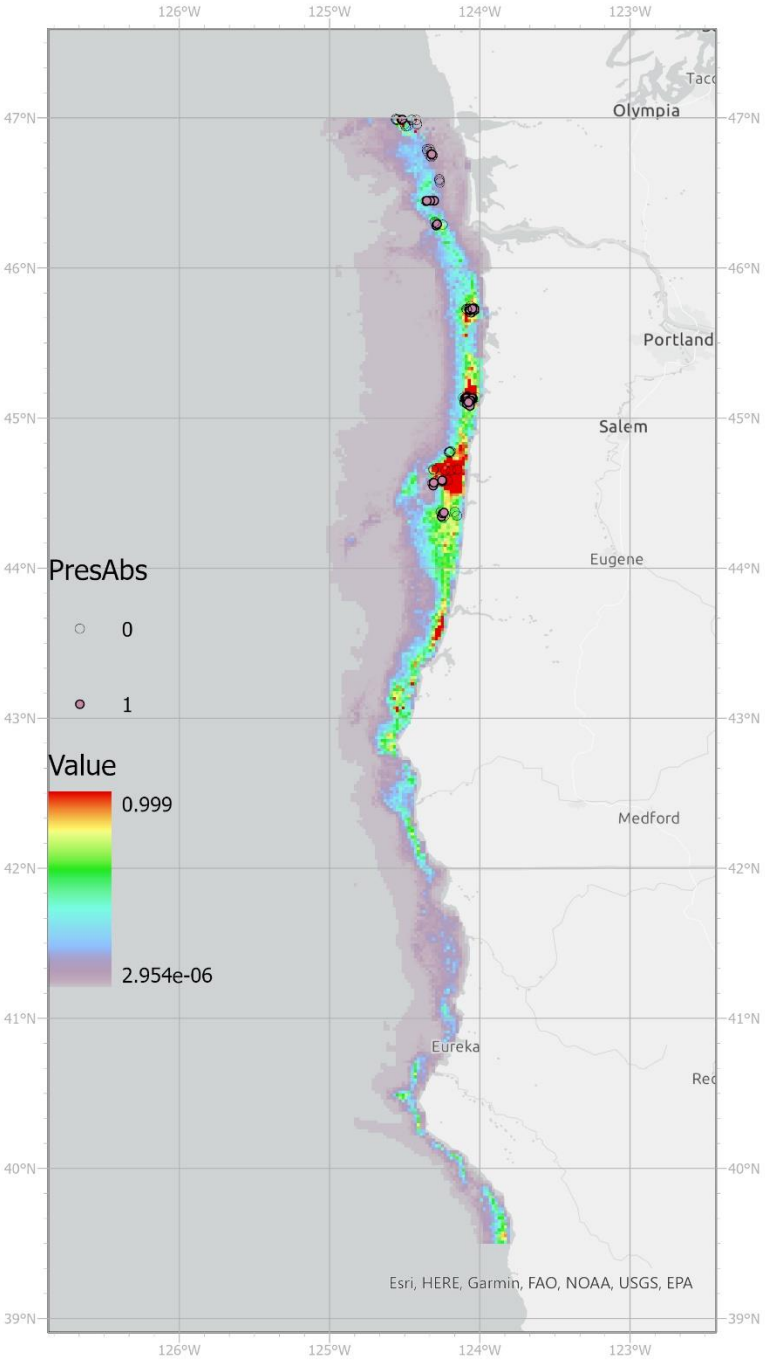
A2. BRT Model with Hemery Layers



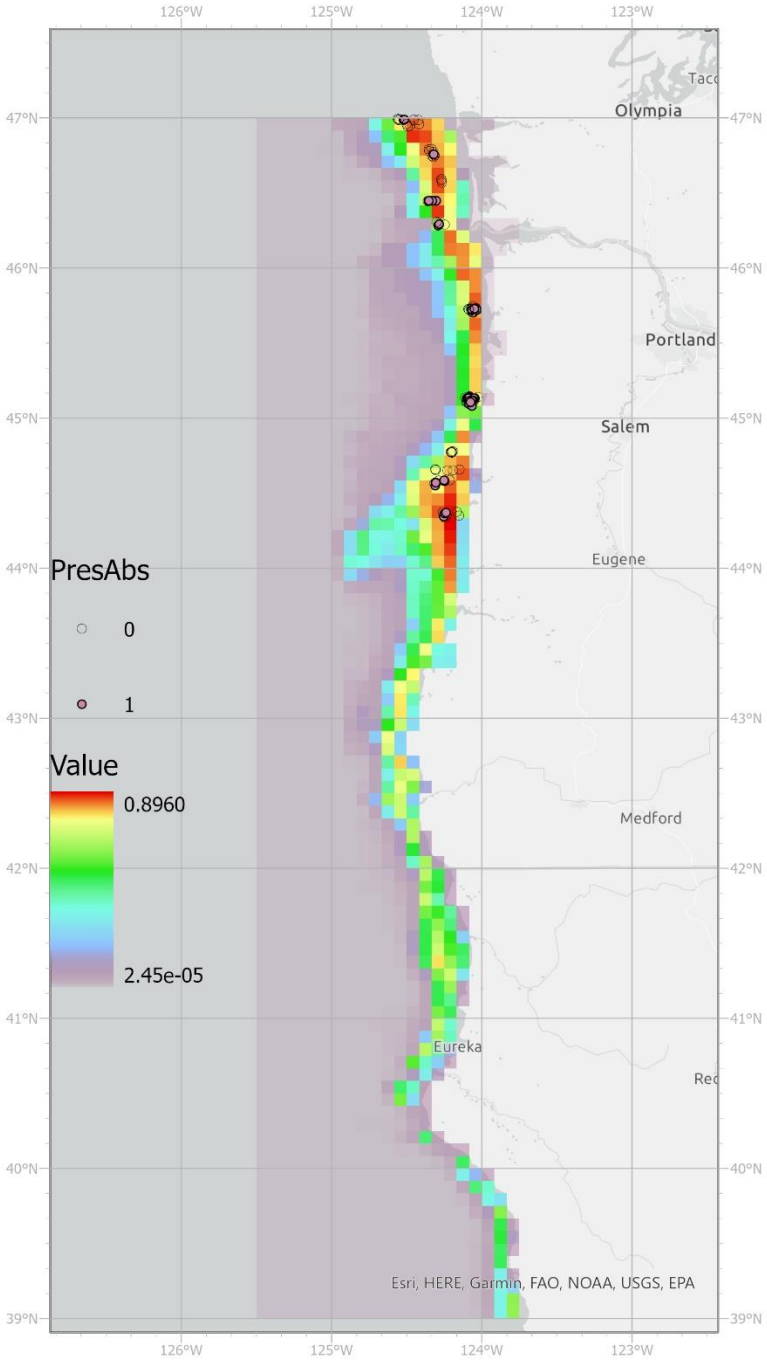
A3. BRT Model with Bio-Oracle Layers



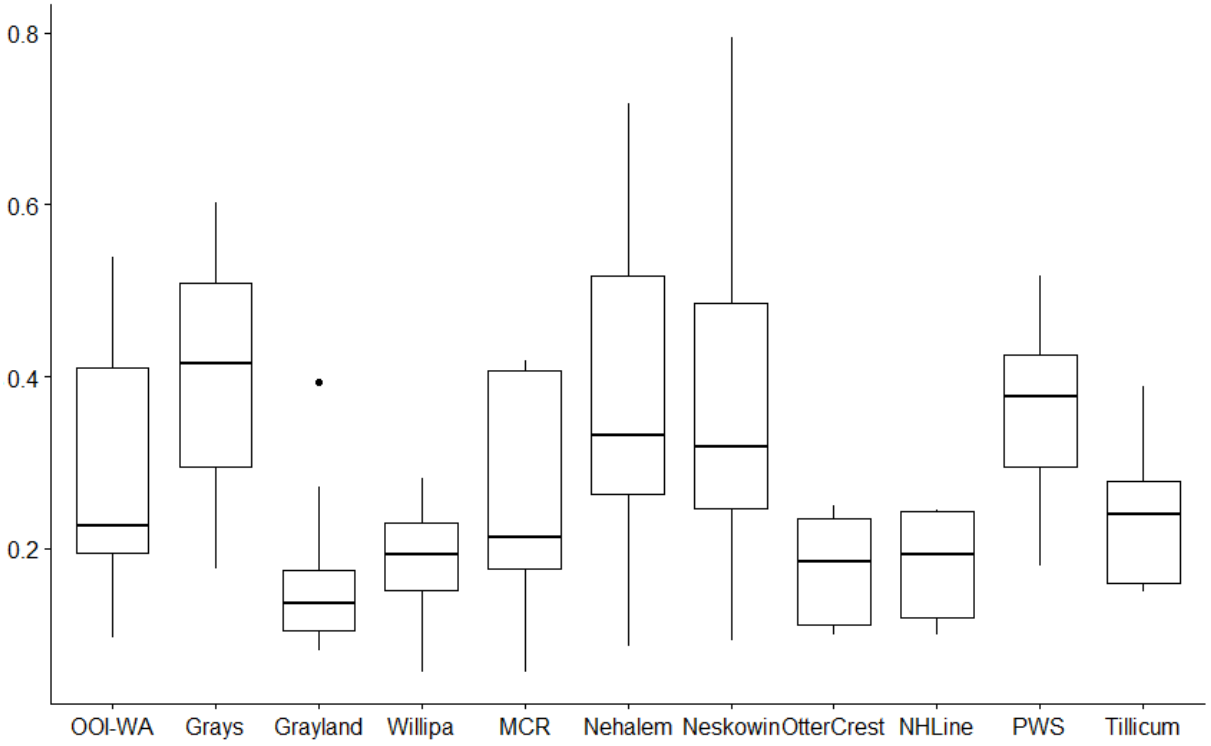
A4. BRT Model with Poti Layers



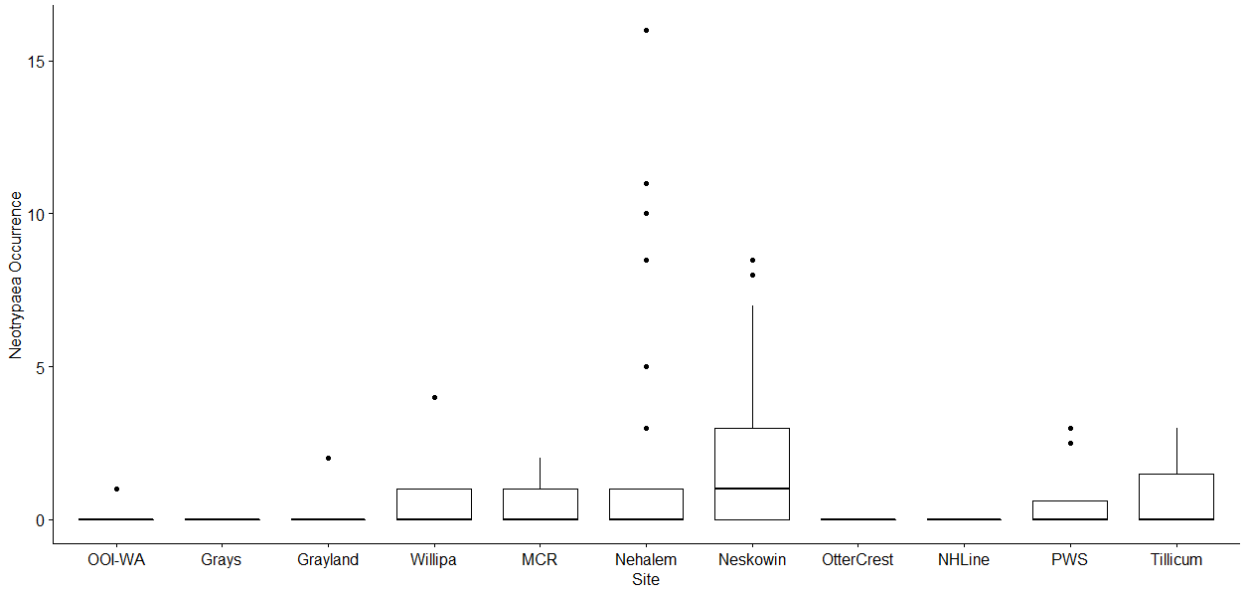
A5. MaxEnt Model with Hemery Layers



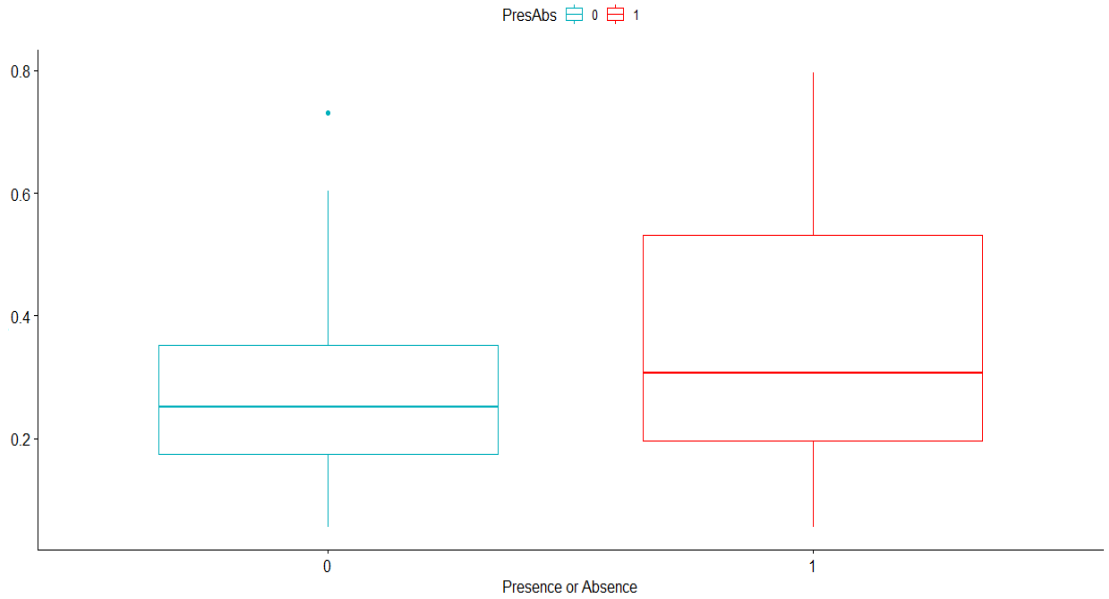
A6. MaxEnt Model with Bio-Oracle Layers



A7*. Probability of Occurrence of the Poti BRT model at each site selected for 2021-2022 sampled sites. The sites were selected based on areas with absences (low probability of occurrence) and presences (high probability of occurrence). Locations with higher means had specific locations with higher probability of presence.. See Table 6 for true absences and true positives of the Poti BRT model.*



A8. *Numbers of* Neotrypaea *per box core grab at each site. Neotrypaea Occurrence is the #/box core. Areas with a high number of Neotrypaea per box core include Nehalem and Neskowin, OR.*



A9. *Box plot of probability of presence of the Poti BRT model at sites with either no* Neotrypaea *or at sites with* Neotrypaea.

Appendix Table 1. *Tukey Post-Hoc for the sediment. Significant values are in bold*

Tukey Post-Hoc						
	p adj					
Difference	*Shewanella*	*Pseudoalteromonas*	*Cellulophaga*	*Bacillus*	*Desulfobulbus*	*Electrothrix*
present-absent	**4.00E-07**	0.569443	0.3642509	0.35633	**5.60E-06**	0.1668385
offshore-estuary	0.1713168	**0.0031213**	**0.0385448**	0.30577	0.9859888	0.2575144
present:estuary-absent:estuary	**0.0021284**	0.9998741	0.6502248	0.84298	**0.0036569**	0.997644
absent:offshore-absent:estuary	0.9206656	0.2353226	0.1780899	0.9371	0.9999998	1
present:offshore-absent:estuary	**0.0000254**	**0.0631633**	0.1502847	0.99998	**0.0031069**	0.2972666
absent:offshore-present:estuary	**0.016084**	0.2644994	0.7886378	0.51398	**0.0045009**	0.9979151
present:offshore-present:estuary	0.5547857	**0.074222**	0.7668663	0.81559	0.9999984	0.3960682
present:offshore-absent:offshore	**0.0003042**	0.9444295	1	0.9449	**0.0038775**	0.3186954