8-14-2014

Bacterial Symbioses and the Innate Immune Response of the Model Host: Euprymna scolopes

Andrew J. Collins
University of Connecticut - Storrs, ajcoll585@gmail.com

Bacterial Symbioses and the Innate Immune Response of the Model Host: *Euprymna scolopes*

Andrew Collins

University of Connecticut, 2014

All animals enter into beneficial relationships with bacteria. The light organ of the Hawaiian Bobtail squid, *Euprymna scolopes*, is a unique model for studying the establishment and maintenance of a symbiosis between a host and a single bacterial species, *Vibrio fischeri*. This bacterium inhabits a specialized structure known as the light organ and provides counter-illumination to mask the silhouette of the predator as it hunts for food during the night. Hemocytes, the primary innate immune cells, preferentially bind and phagocytose non-symbiotic at higher rates than their symbiont, but this can change with the colonization state of the animal. A goal of this work was to use high-throughput sequencing to identify genes expressed within hemocytes of adult animals. Of the many genes identified was a novel peptidoglycan recognition protein, EsPGRP5, which is one of the most abundant transcripts in circulating hemocytes.

In addition to the light organ, female squid have an accessory nidamental gland (ANG) which contributes to making the jelly coat that covers the squid's eggs. This gland is full of epithelium-lined tubules, most of which support a dense population of bacteria. Unlike the light organ however, the bacterial population in the ANG is a consortium of several species, not just one. This work characterized the bacterial symbionts within the ANG and showed that the population is partitioned, with certain tubules containing a single taxon.

The dominant members of the ANG consortium are members of the ubiquitous *Roseobacter* clade, a group of marine bacteria within the *Rhodobacterales*. Several

organisms were cultured from the ANGs of *Euprymna scolopes*. The genomes of these organisms were sequenced and analyzed for secondary metabolites. Among these were siderophores and homoserine lactones.

The results of this work establish a database for future research into the molecular mechanisms of hemocytes. The presence of several genes can lead to the future characterization of many proteins and signaling pathways within hemocytes. This work also established the accessory nidamental gland as new model for consortial symbioses in *E. scolopes*. Comparisons between this new symbiosis and the light organ will shed light on how animals maintain many different bacterial communities.

Bacterial Symbioses and the Innate Immune Response of the Model Host: *Euprymna scolopes*

By

Andrew Jeffrey Collins

B.S. University of Puget Sound, 2007

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

At the University of Connecticut

2014

APPROVAL PAGE

Doctor of Philosophy Dissertation

Bacterial Symbioses and the Innate Immune Response of the Model Host:
Euprymna scolopes

Presented by

Andrew Jeffrey Collins, B.S.

Major Advisor _____
Dr. Spencer Nyholm

Associate Advisor _____
Dr. Joerg Graf

Associate Advisor _____
Dr. Daniel Gage

Examiner _____
Dr. David Benson

Examiner _____
Dr. Steven Geary

University of Connecticut
2014

Acknowledgements

I would first like to thank my advisor, Dr. Spencer Nyholm for supervising my research for the past seven years. I am also grateful to the members of my committee, Dr. Joerg Graf, Dr. Daniel Gage, Dr. David Benson and Dr. Steven Geary, who asked held me to a high standard and asked really good questions.

Past and present members of the Nyholm Lab profoundly affected my development as a scientist. I thank our former post-doctoral fellow Dr. Bethany Rader for our constructive discussions on science. I also acknowledge Tyler Schleicher who was a part of our lab for 6 out of my 7 years at UConn. He was an indispensable resource and partner in running the lab, tending to the animals and reviewing my research. I also mentored many students over the years and their inquisitive nature, work ethic and optimism helped me through many long days.

I thank many people who helped me behind the scenes, including my parents who always supported me and encouraged me particularly when experiments were not working and I had seemingly reached my limit. My siblings, Courtney, Jon, Laura and Beth, for always making holidays and vacations fun and enjoyable. I especially thank my girlfriend Jess, who sacrificed many weekends to come into lab with me as I took care of the squid and finished up experiments.

Finally, I acknowledge my fellow graduate students, who are the soul of research at UConn. It was the people I saw every day, working shoulder-to-shoulder in the lab that provided the greatest wisdom and guidance over the years. I especially need to thank my roommate and my truest friend, Reed Goodwin. Over the five years we lived together, we had many fruitful discussions of experiments, teaching, careers and football.

I am indebted to him for so much advice, guidance and meaningful conversation over beers. Things would have turned out very differently without his support.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Beneficial microbes

Recently, it has been estimated that there are 10^{30} bacteria on the planet (1). Taken together, the mass of carbon in these prokaryotes is nearly equal to that of all the plants on earth combined (1). This "unseen majority" plays a critical role in many essential geochemical processes, including the carbon cycle and nitrogen fixation.

One of the greatest contributions to the earth's ecosystem is the fixation of atmospheric nitrogen as ammonia into the earth's soils. All life on planet earth requires nitrogen to make amino acids, the building blocks of proteins, but most nitrogen exists in as an inert gas in the earth's atmosphere. Bacteria fulfill a key role through natural nitrogen fixation; transforming the inert nitrogen gas into a useable form. Many bacteria form special associations with plants to achieve this. *Sinorhizobium* species form nodules with several leguminous plants while several actinorhizal plants are colonized by *Frankia* species (2,3). After colonizing these plants, these bacteria will convert atmospheric nitrogen to ammonia, which will serve as the plants nitrogen source. Nitrogen from these symbionts and other free-living nitrogen fixing bacteria will disperse through the food web to provide higher organisms with the nitrogen necessary to create amino acids. The earth remains a livable planet for higher organisms due to the nitrogen-fixation performed by these bacteria.

Many bacteria also inhabit the surfaces on and within animals. Most bacteria associated with animals live in the gut and serve beneficial or commensal roles. In each gram of human feces, there are 10^{11} bacterial cells (1,4). These bacteria are helpful to host animals in many ways, from out-competing pathogens to breaking down otherwise

indigestible nutrients. While there are many other microenvironments inhabited by bacteria on and within animals, the concentrations of these populations are a small fraction of what is in the gut. For example, most of the skin only has a few thousand cells per square centimeter (1).

As scientific models however, guts are very complex systems to study, particularly in vertebrates. Due to the many interactions between the host and its microbiota as well as between different members of the microbiota, the impact of any one member in such an environment is difficult to tease apart. In order to understand how a host interacts with a particular bacterium, simpler models are needed.

One way to address this issue is to raise animals with a restricted bacterial consortium, containing one or just a few different members. Mice have been successfully raised in sterile conditions making bacteria undetectable by PCR (5). These mice can then be colonized with a single bacterial species or a consortium from either mice or another animal, including the consortium from a human gut. The gut microbiota implanted into these germ-free animals can induce significant changes in the animal. For example, in mice with no gut microbiota, the gut epithelium stops producing certain fucosylated sugars (6). Colonizing the gut with *Bacteroidetes thetaiotaomicron* restores the production of those sugars. However, a mutant of *B. thetaiotaomicron* that lacks the enzymes to break down those sugars does not (6). Perhaps more startling, when mice were colonized with the microbiota of overweight mice, the infected mice had an increased appetite and gained weight, suggesting that the microbiota could have enormous implications for health and animal behavior (7).

While using animals raised in sterile environments is one way to investigate how animals interact with their microbiota, certain assumptions must be made. For one, changing the microbiota of an animal can lead to developmental changes in the animal (8). This introduces another variable into any experiment that could have confounding results on the experiment. Furthermore, colonizing the animal with one or a few organisms creates an artificial environment that may not reflect true interactions in a very complex gut. Another way to investigate the complicated symbiosis in an animals is to use a simpler model, such as the leech gut and the light organ of *Euprymna scolopes*, the Hawaiian bobtail squid (9-12). Both of these model hosts harbor a natural community of only a few members.

The gut of the medicinal leech is unique. The leech feeds on the blood of vertebrates and has predominantly two symbionts, *Aeromonas veronii* and the *Rickenella*-like organism, *Mucivorans hirudinis* (10, 13, 14). There are many remarkable interactions between these organisms in the leech. An interesting change in the population of the leech gut symbionts occurs in the days after a leech consumes a blood meal. *M. hirudinis* is an obligate anaerobe and usually dominates the microbial population of the leech. However, immediately after feeding, *A. veronii*, a facultative aerobe, grows to match the population of *M. hirudinis*, before decreasing over the next few days, leaving *M. hirudinis* as the dominant symbiont (15). This population transition has been hypothesized to be a result of the interplay between the two symbionts. *A. veronii*, being a facultative aerobe, could thrive on any oxygen available in the blood meal, but once the oxygen is depleted, the anaerobic environment created is ideal for *M.*

hirudinis, which then degrades the mucins provided by the host cells and ferments those products as a source of carbon and energy (13, 15).

The light organ of the Hawaiian bobtail squid, *Euprymna scolopes*, has been a model for animal-microbial interactions for nearly 25 years (11). *E. scolopes* is a nocturnal predator that is endemic to the Hawaiian archipelago. At night, the squid forages for shrimp, but is simultaneously in danger of being eaten by predators. In order to avoid detection, the squid uses the light from the bioluminescent bacterium *Vibrio fischeri* as counter-illumination to mask its silhouette (16). *V. fischeri* lives in the light organ of the squid, where it is the only bacterial inhabitant (11). Unlike the complex communities of guts, this binary symbiosis between *V. fischeri* and *E. scolopes* is a simple model that is ideal for investigating how animals interact with a bacterial symbiont.

Early work in the squid-vibrio symbiosis focused on the morphogenesis of the epithelial tissue on the exterior of the light organ. Within a few hours of exposure to *V. fischeri*, the ciliated appendages around the light organ begin to regress and the epithelial cells undergo apoptosis. This was found to be due be a reaction to the presence of lipopolysaccharide (LPS) and peptidoglycan, components of the cell envelope of Gram negative bacteria (17, 18). This is an irreversible signal and will lead to the complete regression of these outer epithelial appendages in 4 days (19). These outer appendages are critical to the selection of *V. fischeri* from a background of some two million bacterial cells that inhabit a milliliter of coastal seawater. Gram negative bacteria will aggregate in mucus secreted by these appendages (20). Interestingly, the mucus is only secreted in the presence of peptidoglycan. Through mechanisms that are not entirely understood, *V.*

fischeri out-compete other non-symbiotic bacteria within the mucus and migrate to a pore in the light organ surface which leads to the deeper "crypts" where the populations of *V. fischeri* are contained (21).

Once the interior crypt spaces are colonized by *V. fischeri*, the bacteria enter into a "rhythm" with the animal. During the day, when the squid is mostly inactive, the bacterial population in the light organ increases. At night there are 10^8 to 10^9 bacterial cells within the light organ of the adult animal (21). This increase in cell density leads to an increase in expression of luminescence due to quorum sensing; where an increased cell density leads to a higher concentration of an autoinducer molecule which permeates other cells and induces expression of the lux operon (22). Many genes, including luciferase and the genes necessary to make the luciferase substrate, luciferin, are encoded by the lux operon, leading to more light production and brighter cells at night. At dawn, the squid vents approximately 95% of the bacteria within the light organ into the surrounding seawater (23). This seeds the water with symbionts for newly hatched squid that have yet to be colonized and "resets" the diel rhythm of the symbiosis.

Other changes occur during this time however. In addition to the regulation of luminescence, the bacteria undergo a unique switch of metabolism. During the day, the bacteria respire glycerol anaerobically, which is a pH neutral process. However, at night, the bacteria begin to ferment chitin, which will create acid and reduce the pH in the crypt spaces (24). The low pH has secondary effects, such as inactivating the peptidoglycan recognition protein, EsPGRP2, which degrades peptidoglycan monomers (also known as tracheal cytotoxin or TCT) secreted by the *V. fischeri* cells (25). This in turn would

presumably lead to the accumulation of TCT in the central core where it could reach concentrations sufficient to be a signaling molecule or toxin.

Innate immunity

Unlike many vertebrate animals, *E. scolopes* lacks an adaptive immune system. A key feature of the adaptive immune system is the ability to "remember" one microorganism or antigen over time. A classic example of this is the "memory" B-cell, which can become active after interacting with a specific antigen from a prior infection. This B-cell changes into a "plasma" cell which will create antibodies to identify the pathogen for phagocytic cells to destroy (26).

All animals have some kind of innate immunity. This can include macrophage-like hemocytes, the production of antimicrobial peptides or other proteins that inhibit the growth of microbes, such as reactive oxygen species (ROS) or moieties such as siderocalin which disrupts bacterial iron acquisition. All of these factors help to shape the microbiota that inhabit any one animal. One particular example lies in the medicinal leech. Symbiotic *A. veronii* have a type III secretion system (T3SS) which they use to avoid phagocytosis by the host immune cells (14). Non-symbiotic species of *Aeromonas* lack this T3SS and are cleared from the animal.

In *E. scolopes*, reactive oxygen species play a critical role in selection of *V. fischeri* from the surrounding seawater. In the epithelial arms of the light organ, nitric oxide synthase (NOS) synthesizes nitric oxide (NO), which is a common signaling molecule but at high concentrations can also be cytotoxic. The NO synthesized is believed to prevent hyperaggregation of bacteria on the light organ's ciliated field (27). It is only after *V. fischeri* colonizes the light organ that the NO production is attenuated.

In addition to the NO produced at the onset of colonization, the squid produce a halide peroxidase, which produces a powerful ROS, hypohalous acid from hydrogen peroxide (28, 29). *V. fischeri* produces a periplasmic catalase which will reduce the hydrogen peroxide to oxygen and water (30). This deprives the haloperoxidase of its substrate, hydrogen peroxide, and prevents the formation of hypohalous acid.

Another interesting example of innate immunity is the production of proteins that sequester iron. Iron is required for many cell for many cell functions, including respiration, detoxification of ROS and many enzymatic reactions (e.g. aconitase in the TCA cycle). When pathogenic *E. coli* invades the intestine of an animal, the surrounding tissues produce iron chelators such as lactoferrin, siderocalin and transferrin (31). Transferrin is used as an iron carrier which can be used to sequester any free iron that exists outside of the host cells. Another protein produced upon infection is siderocalin, which binds to enterobactin, a siderophore produced by enteric bacteria such as *E. coli* and *Salmonella* species. The siderocalin prevents siderophores from scavenging the host tissue for iron to be used by the infecting bacteria. Production of siderocalin is up-regulated during times of sepsis and inhibits the growth of organisms that are dependent upon siderophores for iron acquisition (32). In this sense, the host attempts to "starve the invaders" by not allowing the pathogen access to the iron necessary for basic cellular reactions. Mice who do not produce siderocalin are more susceptible to bacteremia and death from bacterial sepsis (32).

One of the most interesting aspect of innate immunity in *E. scolopes* are the squid's hemocytes. These macrophage-like cells are the primary immune cell of *E. scolopes*. At the onset of colonization, these cells traffic to the light organ appendages,

increasing their number significantly after only 2-3 hours (33). It is hypothesized that these hemocytes play a role in reorganization of tissue and atrophy associated with tissue remodeling (33). These numbers peak at approximately 24 hours after hatching and the hemocytes remain in there until the appendages completely regress 4 days after hatching.

Aside from hemocyte trafficking to the outer epithelium of the light organ, another remarkable behavior is their ability to preferentially phagocytose non-symbiotic bacteria. When hemocytes are extracted from adults and challenged with the symbiont *V. fischeri*, only a few cells are bound and phagocytosed (34). However, challenging the hemocytes with other closely-related bacteria, like *P. leognathi* or *V. harveyi*, results in many more cells binding to the hemocytes. Such specificity in the behavior of an innate immune cell is surprising given these cells have no clear mechanism for differentiating between similar bacteria. However, if the symbionts of the light organ are removed with antibiotics, this specificity changes. In animals treated with antibiotics, *V. harveyi* binds to hemocytes at a high level, roughly 10 bacteria per hemocyte which is similar to untreated control animals. Over several days of treating animals with antibiotics, binding of *V. fischeri* to hemocytes increases until it reaches a similar rate as *V. harveyi* on the 5^{th} day of treatment (roughly 10 bacteria bound per hemocyte). These data suggest that *V. fischeri* "educates" the host hemocytes to tolerate the symbionts and that the population of hemocytes gradually loses this tolerance after the *V. fischeri* cells in the light organ are removed (34).

The *Roseobacter* clade and symbiosis

The *Roseobacter* clade is a dominant clade of marine bacteria that represents 10% of cells in the open ocean and up to 25% in coastal waters (35). These

Alphaproteobacteria were initially characterized as non-oxygenic photosynthetic heterotrophs and named for their striking pink color, which is caused by the production of the pigment bacteriochlorophyll (36). More recent research has characterized many new members of this clade. While not all produce the classic bacteriochlorophyll pigment, these closely related bacteria are a dominant taxa in seawater and have many important ecological contributions.

One of the most widely studied interactions in the *Roseobacter* clade is its metabolism of sulfur compounds and their contribution to the sulfur cycle (37, 38). Dimethyl sulfide (DMS) is a volatile gas that is the main contribution of sulfur to the atmosphere by the ocean. It has been hypothesized that the release of the DMS from marine environments plays a key role in the creation of clouds in the atmosphere. Once released, DMS reacts in the atmosphere to create aerosols, known as non-sea salt sulfate aerosols, which are believed to be the primary sources of cloud formation at the marine atmospheric boundary (39). One of the primary sources of DMS is the degradation of dimethylsulfoniopropionate (DMSP) by bacterioplankton (37). DMSP is a common metabolite produced by many algal organisms. In coastal environments, it is believed that *Roseobacter* clade organisms are responsible for most of this transformation as they are they are the most commonly isolated organisms that can both hydrolyze DMSP to DMS and utilize DMSP as a carbon and sulfur source (40, 41). Moreover, in culture experiments, 80% of *Roseobacter* clade organisms could convert DMSP to DMS (38). This supports the hypothesis that *Roseobacter* clade organisms play a large role in sulfur cycling in marine environments and could further affect the climate of the ocean from the clouds generated by the DMS released by these organisms.

Most *Roseobacter* organisms have been cultured from seawater samples, but many have also been isolated in association with higher organisms. Many are presumed to be epibionts, which stems from the ability for many bacteria of the *Rhodobacterales*, including *Roseobacter* clade organisms, to be among the dominant early colonizers on surfaces in seawater (42). The first two isolates of the *Roseobacter* clade, *Roseobacter denitrificans* and *Roseobacter litoralis*, were isolated from the surface of seaweed (36). Other examples of *Roseobacter* clade organisms living with other animals include sponges. *Ruegeria sp.* KLH11 was isolated from a sponge and has been shown to have two differing lifestyles, depending on the cell density of the bacterial communities (43). *Phaeobacter sp.* 2.10 has been shown to have a beneficial effect on algal hosts. Originally isolated from *Ulva australis*, this microbe has been shown to produce auxins, which stimulates growth of the algae, and also produces the antimicrobial tropodithietic acid (TDA, 44, 45). TDA has been shown to be beneficial as another closely-related *Roseobacter* clade organism, *Phaeobacter inhibens* DSM2663, produces TDA and has been used as a probiotic in aquaculture systems (46, 47).

However, one fascinating niche of *Roseobacter* clade organisms is the accessory nidamental gland (ANG) which is part of the female reproductive system of some cephalopods. While not all cephalopods have an ANG (octopuses and the colossal squid are notable exceptions), all ANGs studied to date contain bacterial consortiums with *Roseobacter* clade organisms being major constituents (48-51). This suggests a remarkable evolution between these common marine bacteria and higher animals. The ANG is a very structured organ and creates a unique niche for *Roseobacter* clade organisms, but the purpose of the ANG and its bacterial symbionts must be elucidated.

Summary

Euprymna scolopes is a fantastic model for studying the interactions between animals and their microbiota. Not only does it have a binary symbiosis in the light organ, but also a unique association with one of the most abundant lineages in the ocean. The work presented here establishes many new and exciting avenues of research. High throughput transcriptome sequencing of the squid's hemocytes, the primary immune cell in the squid, will allow a better understanding of how these cells differentiate between closely related bacteria. This research is also the beginning of investigating the consortium within the accessory nidamental gland. Characterizing these bacterial symbionts will hopefully lead to determining the function of this organ which is widespread throughout the cephalopods.

Contributions of other researchers

Dr. Bethany Rader and Tyler Schleicher had substantial contributions to the second chapter of this work, including qPCR and proteomics. Barrett Nuttall was instrumental in raising young squid used in Chapter 3. Matthew Fullmer performed the taxonomic analysis of the *Roseobacter* clade genomes in Chapter 5. I am also indebted to the many undergraduates that I have worked with as a member of the Nyholm Lab who fed the squid, helped with many experiments and asked really good questions.

Chapter 2. Hemocyte Transcriptomics and EsPGRP5[1]

Introduction

All animals have co-evolved associations with microorganisms. These symbiotic associations have been a major driving force in eukaryotic evolution, including the vastly influential bacterial endosymbiotic origin of essential organelles as championed by the late Lynn Margulis (52). It is also not surprising that the co-evolved microbiota also influenced the evolution of animal biology. There has been a long-standing view that the immune system evolved to counter infection by pathogens and to exclude microorganisms from animals. However, recent evidence suggests that the immune system also plays an important role as an interface between hosts and their natural co-evolved benign or beneficial microbiota (53-55). This interface involves a "dialog" between host and symbiont(s) that ensures that these associations are established and maintained. Growing evidence suggests that these beneficial microorganisms have influenced the evolution of both the innate and adaptive immune systems (54, 56, 57). Gnotobiotic studies in mammals have demonstrated that proper development of components of the adaptive immune system require the presence of the host's normal microbiota. For example, the development of the gut-associated lymphoid tissue and differentiation of various T-cell classes are greatly influenced by the bacterial communities in the gut (58-60). It has also been proposed that the adaptive immune system itself evolved to deal with an increasingly complex gut microbiota (56).

[1] Portions of this chapter have been published previously. Reference: Collins, AJ, Schleicher TR, Rader BA, Nyholm SV. 2012. Understanding the role of host hemocytes in a squid/vibrio symbiosis using transcriptomics and proteomics. *Front. Immunol.* **3**:91.

Invertebrates, which comprise the vast majority of animals, lack an adaptive immune system and must rely solely on innate immunity to interface with a microbial world. Despite any obvious mechanism to achieve the precision with which the adaptive immune system targets specific pathogens, there are many examples of highly specific associations formed between microorganisms and invertebrates (e.g. insect bacteriocytes, termite hindguts, the trophosomes of hydrothermal vent tubeworms, and the light organ of squids; 21, 61-64). What are the mechanisms for achieving such specificity? Many of these associations exist by creating specialized microenvironments that favor colonization by microorganisms and sequester them to specific cells or tissues (e.g. bacteriocytes, anaerobic, pH or nutrient-specific environments as in the gut; or specialized tissues as in light organs; 21, 65, 66). However, increasing evidence suggests that the innate immune system of animal hosts is also important in maintaining these associations (34, 53, 54, 67, 68).

The light organ symbiosis between the Hawaiian bobtail squid, *Euprymna scolopes*, and the bioluminescent bacterium, *Vibrio fischeri*, is used as a model system to study the mechanisms by which beneficial bacteria colonize animal hosts (Figure 1; 21, 69, 70). The symbionts reside within the epithelium-lined crypt spaces of a specialized light organ (to densities of 10^9 *V. fischeri* cells/adult squid). This organ is directly connected to the environment via ciliated ducts that facilitate the environmental transmission of the symbionts. Juvenile squid must select *V. fischeri* from a background of 10^6 non-symbiotic bacteria/mL of coastal seawater (21). There are many mechanisms by which the partners ensure this specificity and the host's innate immune system is important in this process. For example, innate immunity effectors like reactive oxygen

species, putative complement members, and phagocytic hemocytes, along with the recognition of microbe-associated molecular patterns (MAMPs) by host pattern recognition receptors (PRRs) have all been implicated in playing a role in this association (18, 23, 25, 27, 28, 34, 54, 67, 71-75). Because the symbiosis is binary with a single host and symbiont, it offers the opportunity to ask how such high specificity is established and maintained in the context of interactions with the innate immune system of the host.

Euprymna scolopes, like other cephalopods, has a single type of immune cell, the macrophage-like hemocyte. These hemocytes can traverse between tissues and enter into the crypt spaces where symbionts reside and the hemocytes appear to "sample" these spaces (23). Previous studies of hemocytes from *E. scolopes* have indicated that they behave like macrophages, binding and phagocytosing bacteria (23, 34, 74). Although the crypts of juvenile symbiotic *E. scolopes* may contain hemocytes with engulfed bacterial cells, hemocytes in adult crypts have never been observed with phagocytosed bacteria despite being found in a microenvironment with high densities of *V. fischeri*. These observations suggest that the squid's hemocytes response may change during development of the association. A study of adult hemocytes revealed that these cells could also differentiate between symbiotic and non-symbiotic bacteria (34). *In vitro* binding assays demonstrated that hemocytes differentially bound various species with the *Vibrionaceae*. *V. fischeri* bound poorly to hemocytes from hosts with fully colonized light organs and significantly more to blood cells from hosts whose symbionts had been removed by curing with antibiotics. No significant change was observed for non-symbiotic bacteria. Taken together, these data suggest that hemocytes can differentiate between *V. fischeri* and other closely related members of the *Vibrionaceae* and

colonization alters the hemocytes' ability to recognize the symbiont. The phenomenon of hemocytes changing their binding affinity to *V. fischeri* after colonization may be analogous to vertebrate immune "tolerance", leading to homeostasis (symbiostasis) of the association. This evidence leads us to ask; how can the innate immune system achieve such specificity and how does the symbiont influence the immune response?

In order to better understand the role of host hemocytes in the squid/*Vibrio* symbiosis we have used high-throughput sequencing to characterize the transcriptome and in circulating blood cells of adult hosts with fully colonized light organs. In addition, we have used quantitative RT-PCR (qRT-PCR) to compare expression of several innate immunity genes from both symbiotic and cured hosts, to understand the molecular mechanisms by which the hemocytes response may change due to colonization. We also began the characterization of EsPGRP5, a highly expressed pattern recognition receptor in circulating hemocytes.

Materials and Methods

Animal Collection

Adult animals were collected in shallow sand flats off of Oahu, HI by dip net and were either maintained in the laboratory in re-circulating natural seawater aquaria at the Hawaii Institute of Marine Biology or at the University of Connecticut with Instant Ocean artificial seawater (IO ASW, Aquarium Systems, Blacksburg, VA) at 23°C. All animals were acclimated at least 48 h under laboratory conditions and kept on an approximate 12 h light/12 h dark cycle before sample collection (75).

Hemocyte Collection

Squid hemocytes were isolated from adult *E. scolopes* as previously described (34, 76). Animals were first anesthetized in a 2% solution of ethanol in seawater. Hemolymph was extracted from the cephalic artery, located between the eyes, using a sterile 1-mL syringe with a 28-guage needle. An average of 50-100 µL of hemolymph (~5,000 hemocytes / µL) was obtained per animal using this method. Freshly collected hemocytes were washed and resuspended in Squid Ringer's solution (530 mM NaCl, 25 mM MgCl2, 10 mM CaCl2, 20 mM HEPES, pH = 7.5) and hemocyte concentrations were determined by hemocytometer.

Curing Experiments

To generate naïve hemocytes, adult *E. scolopes* were cured of *V. fischeri*. Adult squid were maintained individually in 5-gallon tanks containing IO ASW. For one set of animals, chloramphenicol and gentamicin was added to the seawater to a final concentration of 20 µg/mL of each. The concentration used in these experiments effectively eliminates *V. fischeri* symbionts from the light organ without compromising the health of the host in any detectable way (17, 19, 20, 33, 34, 77). The animals were transferred daily into fresh IO, either with or without antibiotics for 5 days and samples of hemolymph were collected on day 5. The resulting two sets of hemocytes were designated either "normal" (untreated/symbiotic) or naïve (treated/cured). The effectiveness of the antibiotics was determined by dissecting and homogenizing the central core of the light organ, diluting the homogenate and plating an aliquot of the homogenate on seawater tryptone agar (34).

Transcriptome Sequencing

RNA was extracted from the hemocytes of 11 adult squid using the RNAqueous-Micro kit (Life Technologies, Grand Island, NY) and treated with Turbo DNase (Life Technologies Grand Island, NY). Ribosomal RNA was removed using a eukaryote RiboMinus kit (Life Technologies, Grand Island, NY) and quantified using a Bioanalyzer PicoRNA chip (Agilent, Santa Clara, CA). Two hundred nanograms of treated and pooled RNA (approximately 18 ng from the hemocytes of each of the 11 animals) was used to make a 454 cDNA library according to standard protocol (Roche, Penzberg, Germany). The library was titrated with a SVemPCR kit and sequenced on one-half of a 454 picotiter plate using FLX Titanium sequencing chemistry.

Bioinformatics

454 sequencing reads were assembled using the CLC Genomic Workbench with standard *de novo* assembly parameters (mismatch cost = 2, insertion cost = 3, deletion cost =3, length fraction = 0.5, similarity = 0.8, conflicts resolved by voting, minimum contig length = 200 bases). Contigs were used to query the UniProt and NCBI nr protein databases using blastx with a maximum e-value of 0.01. Significant alignments were further analyzed using SignalP (78), InterProScan (79) and ClustalX (80) using default parameters. The output from the blastx searches were imported into the program BLAST2GO (81) and were used to assign Gene Ontology categories. Level 2 hierarchies were used for comparison between proteomic and transcriptomic sequencing. BLAST results were also used to make annotated protein databases to be searched with MS/MS peptide data, taking the longest open reading frame from any given contig and depositing it into an amino acid database.

Western Blot Analysis

After washing extracted hemocytes in Squid Ringer's solution, cells were lysed in 100µl of radioimmunoprecipitation assay (RIPA) buffer (50 mM TRIS, pH = 7.5, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X 100) with a protease inhibitor cocktail added to 1X working concentration (Sigma Aldrich, St. Louis, MO). Protein was quantified using the Bio-Rad RC/DC protein assay. One microgram of protein was loaded into a 12.5% TGX polyacrylamide gel (Bio-Rad, Hercules, CA) and transferred onto a cellulose membrane using the TurboBlot Semi-dry Transfer apparatus (Bio-Rad, Hercules, CA). Blots were washed in 1x PBS for 15 minutes twice, then incubated with blocking solution (4% non-fat dry milk, 0.05% Tween in 1x PBS) overnight at 4°C. Blots were rinsed twice with PBS-T (0.05% Tween in 1x PBS) then incubated in PBS-T for 5 min at room temperature on a shaker. Membranes were incubated with a polyclonal antibody designed to EsPGRP5 (Manufactured by Genscript, Piscataway, NJ) at 1:1,000 in blocking solution for 1 h with shaking. The membrane was washed, as described above, and then incubated with an HRP-conjugated secondary antibody diluted to 1:5,000 for 1 h with shaking. The membrane was washed 3x with PBS-T, then 2x with PBS for 5 min with shaking. Blots were developed with GE Healthcare ECL Western blot Detection Reagent (GE Healthcare, Pittsburgh, PA) and photographed with a CCD camera (Protein Simple, Santa Clara, CA).

Immunocytochemistry

After hemocytes were extracted, the cells were resuspended in 1 mL of Squid Ringer's solution. Circular coverslips were placed at the bottom of a 12-well cell culture dish and covered with 900 µL of Squid Ringer's solution. 100 µl of the resuspended

hemocytes were added to 10 wells and allowed to adhere to the coverslips for 20 min. To challenge hemocytes with bacteria, either *V. fischeri* or *V. harveyi* were added at 50 bacteria per hemocyte and allowed to incubate with the hemocytes for one hour before fixation.

Cells were fixed with 4% paraformaldehyde in 1X Squid Ringer's solution overnight at 4°C. Cells were permeabolized with 1% Triton and washed with 1X mPBS. Primary antibody to EsPGRP5 was added at 1:1,000 and incubated overnight. Cells were washed and secondary antibody was added at 1:10,000 and incubated for overnight. Cells were counterstained with Rhodamine-conjugated phalloidin (an actin stain) and Draq5 (a DNA stain). Micrographs were taken using a Nikon A1R confocal microscope.

Polymerase Chain Reaction

Total RNA from adult symbiotic or naïve hemocytes was extracted using TRIzol reagent (Life Technologies) according to the manufacturer's instructions. Approximately 800 ng of RNA per animal was isolated. Subsequent PCR and qRT-PCR was performed using cDNA template created from 200 ng of total RNA using the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA) and using the manufacturer's instructions for random primers. End-point PCR was performed on a MyCycler Thermal Cycler (Bio-Rad, Hercules, CA) using the following programs: 3 min at 94°C, 30 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 45 s; with a final extension of 72°C for 10 min. Products were separated on a 1% gel, stained with SYBR safe DNA Gel Stain (Invitrogen, Life Technologies, Grand Island, NY) and visualized using a Molecular Image Gel Doc (Bio-Rad, Hercules, CA).

qRT-PCR was then performed in 20µL reactions using gene-specific primers and Evagreen Supermix (Bio-Rad, Hercules, CA). Quantitation was done using an iQ Multicolor Real-Time Detection System (Bio-Rad, Hercules, CA). Amplification was performed under the following conditions: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 55°C for 15 s and 72°C for 30 s, followed by a melt curve program, performed to confirm the presence of one optimal dissociation temperature for the resulting amplicons. Standard curves were created using a five-fold cDNA dilution series with each primer set. All efficiencies were between 90% and 105%. Each gene was assayed in three technical triplicates for each of the three biological replicates, and data was analyzed using the ΔΔCT relative quantification method with the 40S and L21 ribosomal subunits as control genes. Primers for both end-point and RT-PCR are seen in Table 2.

Results

We analyzed circulating adult hemocytes by high-throughput 454 transcriptomic and LC-MS/MS proteomic analyses. 454 high-throughput sequencing produced 650,686 reads totaling 279.9 Mb (Table 1) while LC-MS/MS analyses of circulating hemocytes putatively identified 702 unique proteins. Analysis of gene ontologies from these datasets revealed that both the transcriptome and proteome yielded similar types of genes/proteins involved with biological processes, molecular function and cell components (Figure 2).

Despite a large number of reads associated with ribosomal RNA, assembly of the transcriptome resulted in 5,586 assembled contigs with an N50 of 521 bp. Of these, 1,995 were annotated by a blastx search. Many of these were predicted to be involved with cytoskeletal structure. Among innate immunity-related genes, several members,

and/or putative regulators of the NF-κB signaling pathway were identified including TRAF4, TRAF6, REL/NF-κB, inhibitor of NF-κB, NF-κB activating protein, both calcineurin subunits, a regulator of calcineurin and MyD88 (Table 3). Several PRRs were identified including galectin and a previously described peptidoglycan recognition protein (EsPGRP4). An undescribed PGRP, most closely related to PGRP-SC2 from the rotifer *Brachionus majavacas* was also found. Further analysis revealed a complete open reading frame (ORF) consisting of a 759bp transcript.

However, difficulty arose in determining the start codon for this protein, as there are 4 methionine residues encoded in the transcript that could serve as a start codons of a functional PGRP (Figure 3). Any other methionine would produce a PGRP with a truncated and likely non-functional peptidoglycan binding site.

The ribosomal binding sequence (or Kozak sequence) remains relatively undefined in cephalopods, so Western blotting was needed to determine the size of the protein and conclusively determine the proper start codon. Western Blotting with a known size marker revealed a single band at approximately 21kD (Figure 4), which corresponds to the methionine at position 79. The resulting translated protein is 183 amino acids long with a calculated weight of 21.2kD. No signal peptides or transmembrane domains are detected by SignalP or InterProScan. However, it does have four conserved residues that are critical in chelating a zinc cofactor for T7 lysozyme-like amidase activity (Figure 5; 82, 83).

When compared to other peptidoglycan recognition proteins, EsPGPR5 is unique. A neighbor-joining phylogenetic tree created using an alignment from ClustalX shows EsPGRP5 is the most divergent PGRP yet found in *E. scolopes* (Figure 6). With a

predicted pI of 9.6, it is a very basic protein. Moreover, transcriptome coverage shows that EsPGRP5 is one of the most highly expressed genes in circulating hemocytes, placing it in the top 5% of most abundant transcripts (Table 4).

Immunocytochemistry of EsPGRP5 in hemocytes showed that it is located in vacuoles within hemocytes (Figure 7). These vacuoles seem to co-localize with the actin cytoskeleton of these cells. Neither curing the animals of their symbionts nor challenging the hemocytes *in vitro* with *V. fischeri* or *V. harveyi* altered the localization of EsPGRP5 in hemocytes (Figure7B-D). However, curing the animals of symbionts did decrease the expression of EsPGRP5, lowering its expression nearly 5-fold (Figure 8).

Discussion

This study represents the first transcriptomic analyses of the circulating hemocytes from *E. scolopes*. As might be expected, in addition to factors involved with general biological processes, hemocytes express a suite of innate immunity genes under a normal symbiotic state. Our data have revealed a number of genes and protein that have previously been shown to be important in the squid/*Vibrio* association. In addition we have described a fifth PGRP (EsPGRP5), which was among the most abundant transcripts in hemocytes. We also identified novel genes/proteins, for example a putative cephalotoxin and plancitoxin that have not been previously described in invertebrate hemocytes. Finally, we demonstrate that some of these genes are differentially expressed in hemocytes from symbiotic and cured hosts, suggesting that colonization influences gene expression of the cellular innate immune system of *E. scolopes*. The cellular localization and function of many of these genes and proteins have yet to be characterized

in detail but we discuss below the putative function of some of these innate immunity effectors in the context of host hemocytes.

Among the transcripts and identified peptides, many were predicted to be involved in cytoskeletal structure (Table 3). This is not surprising considering that these macrophage-like hemocytes can switch within minutes from a round spheroid to an activated cell with amoeboid characteristics. These activated cells are then capable of migrating through tissues and have been shown to enter the crypts of the light organ where they can interact with *V. fischeri* (23, 74).

Among those genes involved with innate immunity, several putative PRRs were detected including a galectin and two PGRPs (EsPGRP4 and EsPGRP5). Galectins are proteins that can bind carbohydrates and have been implicated as PRRs during the immune response (84). In other mollusks, galectins have been shown to bind microorganisms and are up-regulated in the presence of bacteria (85-87). The role of galectin in this symbiosis has yet to be characterized, but efforts are underway to compare gene expression and protein levels in hemocytes of colonized and uncolonized hosts.

Whether an association is pathogenic or beneficial the immune system must first recognize the microorganism(s). These encounters are mediated by the initial recognition of MAMPs by PRRs (18, 73, 88). PGRPs are one of a suite of PRRs that recognize MAMPs and are evolutionarily conserved (55). Five PGRPs have now been described from *E. scolopes* and prior work has demonstrated that some of these are important during the onset of the symbiosis (25, 72, 89). Analyses of EsPGRP1 and EsPGRP2 protein expression in the juvenile light organ have shown that they have different cellular

localizations within the crypt epithelia and nascent light organ tissues. *V. fischeri* can influence this expression with specific MAMPs. For example, tracheal cytotoxin (TCT) a derivative of peptidoglycan that along with lipopolysaccharide signals early development of the host light organ (17, 18), influences the nuclear loss of EsPGRP1 in host epithelial cells (72). EsPGRP2 which is secreted into the light organ crypts has TCT-degrading amidase activity and may ameliorate the effect of this toxin, especially during times when the symbiont numbers, and presumably TCT concentrations, are high. The ability to have a diverse repertoire of PGRPs may give the host greater ability to respond to peptidoglycan and TCT depending on the MAMP microenvironment that it encounters.

The results of this study describe a fifth PGRP and suggest that EsPGRP5 is one of the most abundant transcripts expressed in host hemocytes. Furthermore, qRT-PCR showed that this gene is down-regulated 4.7 fold in the hemocytes from cured hosts suggesting that colonization of the light organ may directly influence PRRs in hemocytes.

As EsPGRP5 has no transmembrane domain and is not observed as a membrane protein, it seems unlikely to act as a receptor protein. It may be secreted and degrade peptidoglycan like its counterpart, PGRP2. EsPGRP5 has conserved residues for T7-like amidase activity and therefore may play a role in decreasing the toxic effects of peptidoglycan and its monomer, TCT.

Being the most divergent PGRP yet described in *E.* scolopes may be due to the fact that it is from a different cell type as the other PGRPs, which were sequenced from light organ tissue. Hemocytes are a systemic tissue and therefore a PRR such as EsPGRP5 must be able to interact with many different MAMPS from a variety of microorganisms and may not be as specialized as the previously described proteins

characterized from tissue that closely interact with the light organ symbiont *V. fischeri*. This could also explain why curing the light organ of symbionts did not affect the localization of EsPGRP5, as its cellular localization may be more strongly influenced by microbes in the gut or elsewhere in the animal. As EsPGRP5 is the most highly expressed PRR yet found in hemocytes, its role in immunity should be ascertained.

The Toll/NF-κB pathway is an evolutionarily conserved signaling pathway often involved in regulating innate immune genes (90). In *E. scolopes*, many homologs of this pathway have been identified (89) and microarray data demonstrated different expression levels in colonized vs. uncolonized juvenile light organs (91). Our transcriptomic data revealed several members of this pathway, including MyD88, the universal adaptor protein for the TLRs as well as other IL-1 receptor proteins that play an important and well-studied role in inflammation and host defense (92). MyD88 is also involved with interactions of epithelial surfaces with microbial communities. For example in mice, production of certain antimicrobial peptides by Paneth cells is dependent on MyD88, and is sufficient to limit bacterial translocation to the mesenteric lymph nodes without altering the number of luminal bacteria (93). Two TNF receptor-associated factors TRAF4 and TRAF6 were also identified. TRAF4 is a signal transducer of the Toll/IL-1 family. In mouse macrophages, TRAF4 has been shown to negatively regulate the NF-κB signaling cascade in response to the peptidoglycan byproduct muramyl dipeptide (94) while TRAF6 is involved with regulating ROS via the NF-κB pathway in mammalian macrophages (95). Both the A and B subunits of calcineurin were also detected in the transcriptome. Calcineurin is known for modulating immune responses in macrophages and this protein can inhibit both IL-12 and TNFα production. Furthermore, inhibiting

calcineurin *in vitro* increases expression of effector genes, including NOS and NF-κB as much as eight-fold (96). Calcineurin may also be involved in responses to LPS since mice lacking a subunit of calcineurin show tolerance to LPS, with much lower TNFα production than wild type animals (97). Recent transcriptome profiling of other molluscan hemocytes from an oyster and mussel challenged with either pathogenic vibrios or lipopolysaccharide also revealed a number of expressed pattern recognition and signaling molecules related to the Toll/NF-κB pathway, thus highlighting the importance of this evolutionarily conserved pathway in this cell type (98, 99).

Recent analyses of the squid/vibrio system have identified a number of putative members of the complement pathway (71, 75). C3 is the central protein to all three pathways of the complement system and is involved in opsonization and phagocytosis (100). A C3 ortholog was recently discovered in *E. scolopes* and shown be located at the apical surfaces of crypt epithelial cells (71). In this study C3 was found in the hemocyte proteome, but was absent in the transcriptome. End-point RT-PCR showed C3 expression only in hemocytes from cured hosts, explaining its absence from the transcriptome as RNA was isolated from hemocytes extracted from animals still colonized with *V. fischeri*. We do not yet know the turnover rate of the EsC3 protein hemocytes but removing the symbionts led to an increase in gene expression of this main component of the complement pathway. Also identified were two TEPs. TEPs are similar complement-like components and alpha macrogobulins, which have basic functions as proteases and protease inhibitors respectively (101). Among invertebrates, TEPs play an important role in the innate immune response as members of a complement-like system (101, 102). A recent analysis of the light organ proteome found

several TEPs (75). Also identified was an alpha macroglobulin belonging to a family of protease inhibitors and has been shown to be an important component of the innate immune response of arthropods (103).

Among the other proteins identified was a matrix metalloproteinase similar to vitronectin. Vitronectin has been described as a serum glycoprotein involved in regulating innate immunity homeostasis (104). In humans, vitronectin inhibits the membrane attack complex of the complement pathway. Bacteria, including *Haemophilus influenzae* and *Neisseria meningitides* can interact with vitronectin and evade the complement response (105). Recent evidence also suggests that the parasite *Plasmodium falciparum* may use vitronectin-like domains to evade the immune system (106). LC-MS/MS analyses identified a putative vitronectin in circulating hemocytes and results of a qRT-PCR showed that MMP/vitronectin was down-regulated 5.6 fold in the absence of *V. fischeri*. The putative MMP/vitronectin in this study contains a hemopexin-like domain but whether it serves a role similar to vitronectin in the squid host remains to be determined.

Both the transcriptome and proteome identified a protein with high similarity to cephalotoxin (Table 3, 107). In octopus and cuttlefish, cephalotoxins are normally expressed in the salivary glands and are involved with immobilizing prey. This is the first report of this protein in another tissue type. A search of the UniProt database identified a thrombospondin domain similar to ones described in properdin, a proenzyme in the complement cascade (108, 109). Whether cephalotoxin functions as a toxin, a protease or in another capacity in the hemocytes remains to be determined but qRT-PCR showed a 7.1-fold decrease in expression in naïve hemocytes, suggesting that

colonization also influences expression of this gene. A second toxin was also identified with similarity to plancitoxins found in the crown-of-thorns starfish *Acanthaster planci* (110). These proteins are predicted to be deoxyribonucleases and future studies should focus on characterizing any potential DNase activity in squid hemocytes.

Reactive oxygen species are common effectors used by the innate immune system to combat pathogens. In symbiotic associations, ROS have also been shown to be important, often in contributing to microenvironments that may promote a microbial member (27, 28, 111, 112). The enzyme nitric oxide synthase (NOS) is required for generating antimicrobial reactive nitrogen species (RNS) such as nitric oxide (NO). Both NOS and NO are active during colonization of the host and NO was detected in vesicles found within mucus secretions that interact with *V. fischeri* during initiation of the association (27). Colonization leads to attenuation of NOS and NO in the crypts suggesting that the RNS response may change in order to accommodate the symbiont. (27) Although not detected in the transcriptome or proteome, we decided to analyze NOS expression in symbiotic and naïve because of its likely importance in the host immune response. Naïve hemocytes showed a 16.8-fold increase in expression suggesting that colonization leads to a suppression of NOS in hemocytes, similar to what has been described for juvenile light organs (Figure 8).

Another condition that may influence hemocyte gene expression is a changing symbiont population over a 24-h day/night cycle. During the evening when the host is active, there is a full complement of *V. fischeri*, but at dawn, the host expels 95% of these symbionts into the surrounding seawater (23, 113). We do not yet know how gene expression of host hemocytes might change during the day/night cycle and the context of

the microenvironment encountered in the crypt spaces of the light organ, but prior studies have reflected changes in host and symbiont transcription in adult light organ tissues over a 24-h period (24). Light organ gene expression is also altered by colonization in juvenile hosts (91). Our transcriptome was derived from circulating hemocytes from a variety of time points to understand the overall background gene expression in colonized adults. Growing evidence suggests that the immune response fluctuates on a daily rhythm (114, 115) and future studies should also consider how innate immunity effectors and pathways might vary over the day/night cycle in the squid/vibrio association.

We still have much to understand concerning the role of host hemocytes in the squid/vibrio symbiosis, but future research into this cell type will likely yield interesting results. Macrophage-like phagocytes serve at the interface between hosts and symbionts. Besides combating pathogens, invertebrate hemocytes have recently been found to play important roles in mediating bacterial symbioses as in the medicinal leech (14), linked to a primed immune response to specific bacteria in *Drosophila* (116) and promoting symbiont tolerance in the squid/vibrio association (34). My work established a transcriptome database that can be used to further characterize the proteins expressed in hemocytes. Future studies will use quantitative transcriptomic and proteomic methods to compare gene and protein expression between symbiotic and naïve hemocytes and those challenged with *V. fischeri* and non-symbiotic bacteria. From the transcripts presented here, full-length protein sequences can be obtained for either recombinant protein expression or antibody design for immunocytochemistry. The transcriptomic data in this study lays the groundwork for exploring the molecular mechanisms of hemocyte/symbiont specificity in the squid/vibrio symbiosis and adds further evidence

that interactions between beneficial bacteria and the innate immune system are important in mediating those interactions.

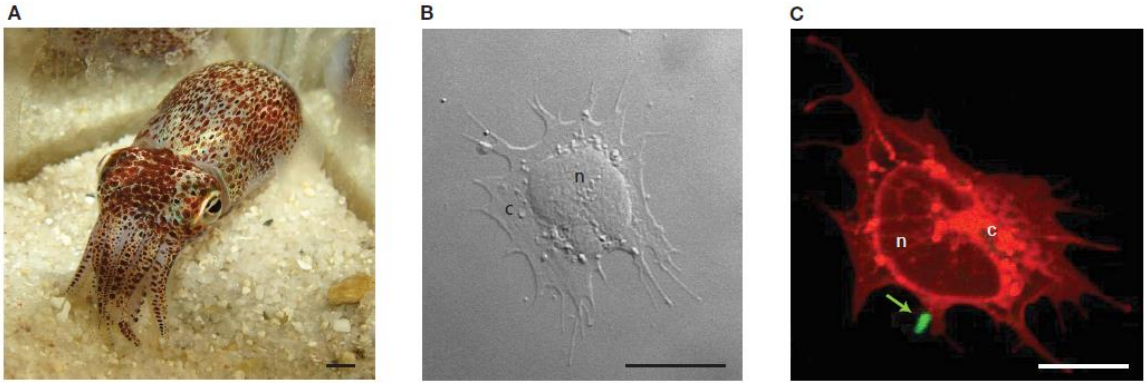

Figure 1. The Hawaiian bobtail squid, *Euprymna scolopes*, and its hemocytes. (A) Adult *E. scolopes*. Scale, 3mm. (B) Hemocyte of the first host as viewed by DIC imaging. The macrophage-like hemocyte is the only circulating blood cell type found in the host. These cells can infiltrate the crypt spaces of the light organ where they can interact with the symbiont *V. fischeri.* Scale, 10µm. (C) Host hemocytes stained with CellTracker (red) viewed by confocal microscopy. A sole *V. fischeri cell* (green, arrow) can be seen bound to the hemocyte. Scale, 10µm. n, nucleus, c, cytoplasm.

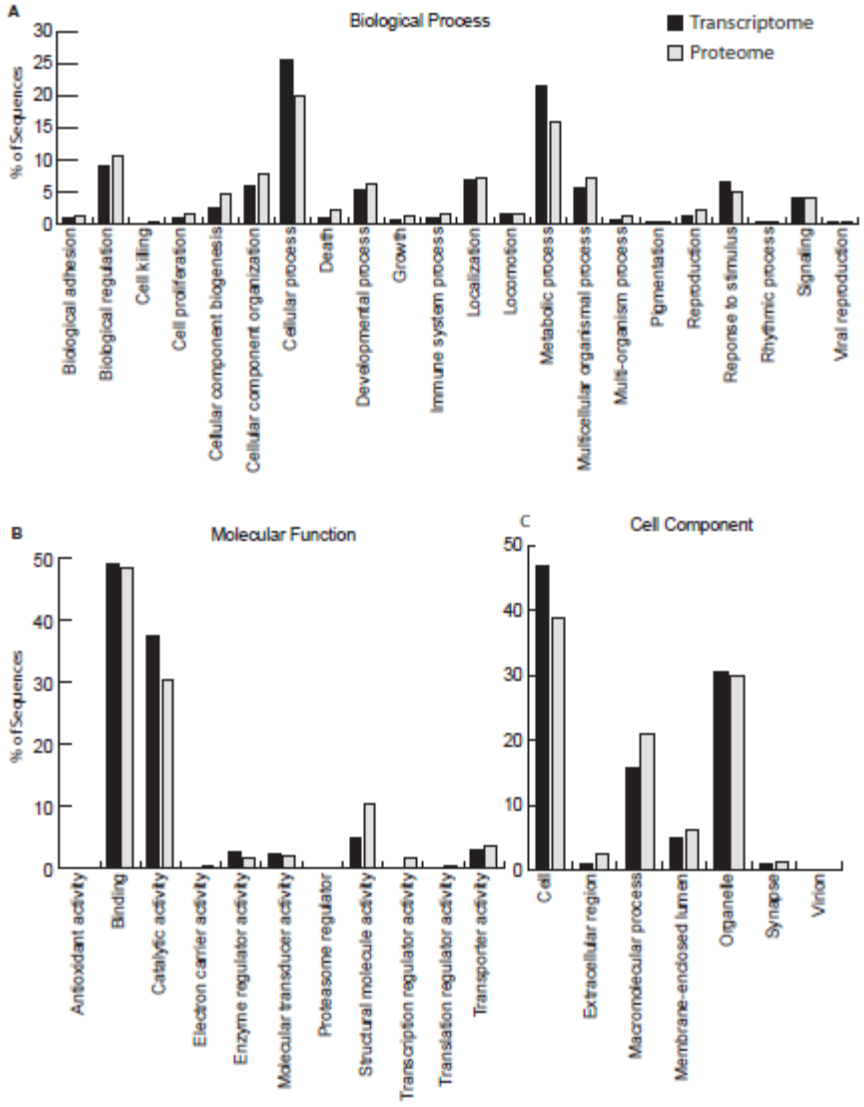

Figure 2. Gene ontology (GO) analysis of the hemocyte transcriptome and proteome. Annotation by GO revealed similar proportions of GO terms in both the transcriptome and proteome (Datasets 1 and 2). Sequences were annotated using previously described BLAST parameters (see Materials and Methods), and level 2 GO terms in **(A)** biological process, **(B)** molecular function and **(C).**

Table 1. Summary of hemocyte 454 transcriptome sequencing

Number of reads	650, 686
Avg. read length	429.9 bases
Total bases	279.9 Mb
No. of rRNA reads (% of total)	567, 545 (87.2%)
Contigs assembled	5, 586
Contigs annotated by blastx	1, 995
N50	521 bases
Unassembled reads	364, 383

Table 2. Primer sequences used in end-point qRT-PCR

Primer name	Primer sequences
Es40sF	AATCTCGGCGTCCTTGAGAA
Es40sR	GCATCAATTGCACGACGAGT
EsL21F	GCCTTGGCTTGAGCCTTCAACTTT
Es21R	GGTGATCGTCAACAAACGCGTGAA
EsC3F	TGCTGTTCCGTTCTGTGAGCACTA
EsC3R	GCAACACACTCTCTTTGAGCGCAT
EsNosF	TGTTCGAGAGTGTTGCGGACATCA
EsNosR	TCACTTGAATGTTGCCGTCTCCAG
EsPGRP5F	TTGGAGCCCACACCAAGTTCTACA
EsPGRP5R	CAAGACAGCAGGCGTTTCAATGCT
EsCephalotoxinF1	TACCACGCGAATTGACACAT
EsCephalotoxinR1	TACTGGAACGCACAAGTTGC
EsVitronectinF1	TGCGCCAAAATACATCAGAG
EsVitronectinR1	TGGGAATATCCCACGGTAAA

Table 3. Innate immunity-related transcripts from 454 sequencing of *E. scolopes* hemocytes

Putative ID	Accession #	Organism	e-value	Putative function
NF-κB Cascade				
TRAF6	AAY27978.1	*E. scolopes*	4e-48	NF-κB signaling
TRAF4	ABN04153.1	*B. belcheri*	9e-50	NF-κB signaling
REL/ NF-κB	AAY27981.1	*E. scolopes*	9e-65	NF-κB signaling
Inhibitor of NF-κB	AAY23980.1	*E. scolopes*	4e-4	NF-κB signaling
NF-κB activating protein	CAX69433.1	*S. japonicum*	1e-36	NF-κB signaling
Calcineurin subunit A	ABO26624.1	*H. discus*	5e-54	Suppressor of NF-κB
Calcineurin subunit B	ACI96107.1	*P. fucata*	2e-56	Suppressor of NF-κB
Regulator of calcineurin	NP_001011215.1	*X. tropicalis*	4e-11	Regulator of calcineurin
MyD88	BAC76897.1	*C. porcellus*	1e-20	Toll-signaling
MAMP Binding				
Galectin	ACS72241.1	*A. irradians*	3e-43	Glycan binding
PGRP-SC2 (EsPGRP5)	ACV67267.1	*B. majavacas*	5e-38	PGN binding/cleavage
PGRP4	AAY27976.1	*E. scolopes*	1e-52	PGN binding
Other				
IK cytokine	ABV24915.1	*C. gigas*	7e-47	Anti-interferon cytokine
CD63 antigen	ABY87409.1	*H. diversicolor*	3e-25	Phagosome protein
Plancitoxin	Q75WF2.1	*A. planci*	1e-68	DNA-degrading protein
MMP/vitronectin	BAC66058.1	*S. kowalevskii*	7e-7	Immune homeostasis
Cephalotoxin	B2DCR8.1	*S. esculenta*	2e-118	Toxin

E. scolopes, *Euprymna scolopes*; *B. belcheri*, *Branchiostoma belcheri*; *S. japonicum*, *Schistosoma japonicum*; *H. discus*, *Haliotis discus*; *P. fucata*, *Pinctada fucata*; *X. tropicalis*, *Xenopus tropicalis*; *C. porcellus*, *Cavia porcellus*; *A. irradians*, *Argopecten irradians*; *B. majavacas*, *Brachionus majavacas*; *C. gigas*, *Crassostrea gigas*; *H. diversicolor*, *Haliotis diversicolor*; *A. planci*, *Acanthaster planci*, *S. kowalevskii*, *Saccoglossus kowalevskii*; *S. esculenta*, *Sepia esculenta*; PGN, peptidoglycan; MMP, matrix metalloproteinase.

KRVYSQYN**M**KVLQI**M**VLCLCLHSLSGATAPKICACALKPTLLINFASKDTC**M**IWPGAC
LEVSHLNSPSWVHVVRNGKI**M**QGWRKSFQIKLCQHQQSCPKIITRKEWGALKPKARTP
LSTPVKYAVIHHSDTPKCHSKMKCIERVRSIQEYHMHHNHWSDIGYNFLIGSDGNVYE
GRGSDTVGAHTKFYNSQSIGICVIGNYSSSRPNWPSLIALKRLLSCLKNNKKLKNDYS
LKGHRDLSPTKCPGKYLYNNITHWPHYKY

Methionine start	Predicted protein mass
Met 9	28.7 kD
Met 15	28.0 kD
Met 52	24.2 kD
Met 79	21.2 kD

Figure 3. Open reading frame of EsPGRP5 with possible start sites. When translated from the EsPGRP5 transcript, there are 4 possible methionine start sites that would produce a protein with a functional peptidoglycan recognition/binding domain.

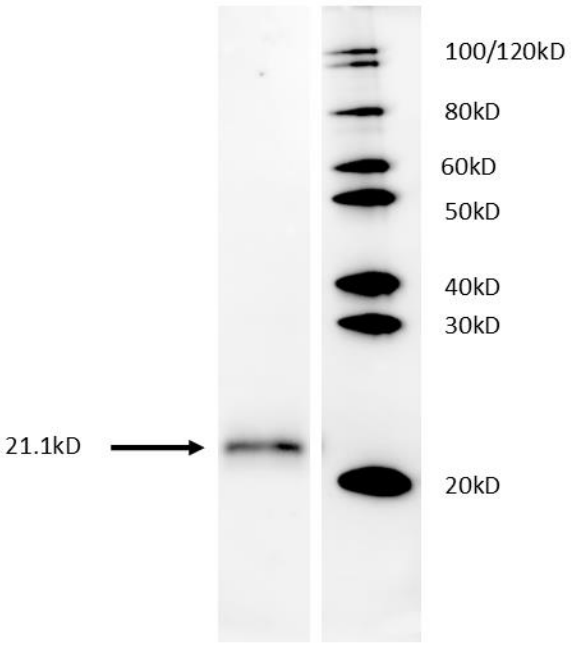

Figure 4. Western blot of EsPGRP5 from hemocytes. A single band detected with an anti-EsPGRP5 antibody shows that one dominant isoform of the protein with a mass of approximately 21.1kD exists in circulating hemocytes. This mass corresponds to one of the possible start codons identified earlier.

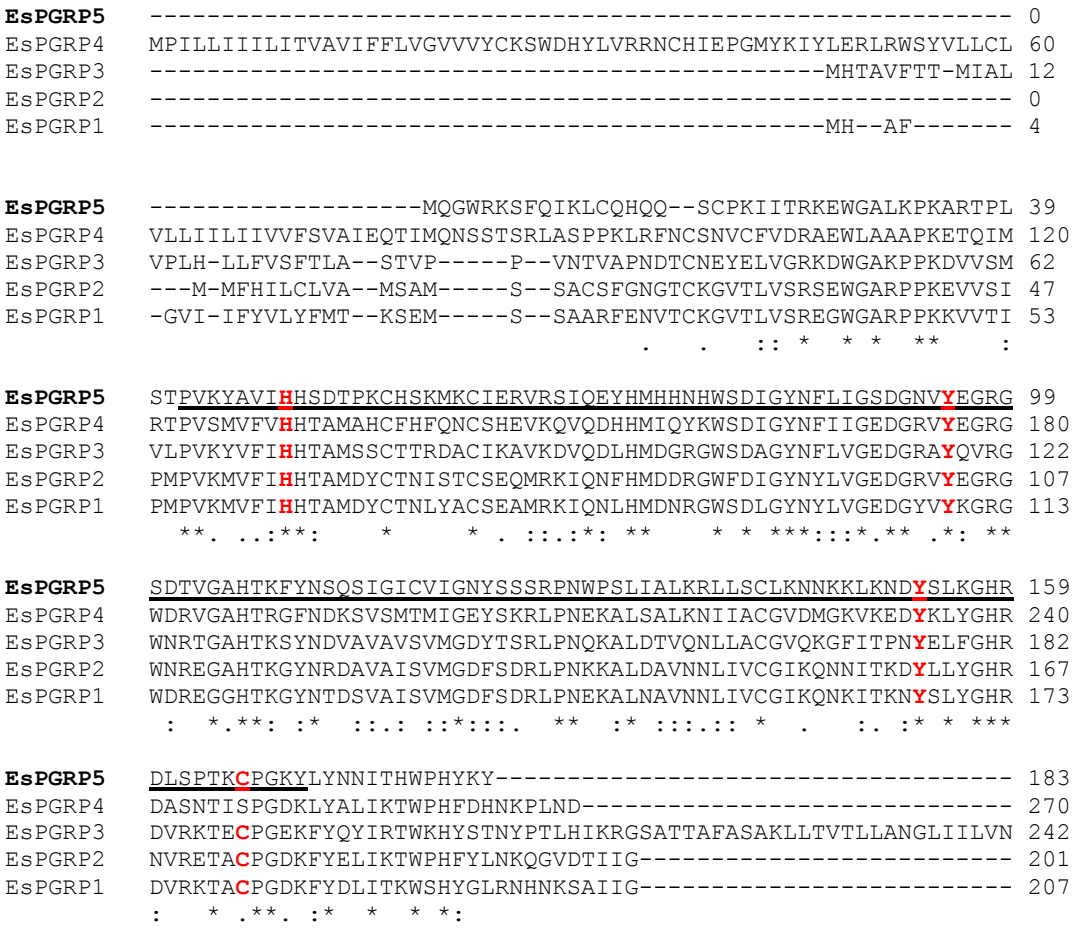

Figure 5. Alignment of *E. scolopes* peptidoglycan recognition proteins. Alignment of EsPGRP5 with the previously described EsPGRPs 1-4. A 183 residue protein was found from transcriptomic sequencing with four conserved residues that indicate potential amidase activity (red). The peptidoglycan binding domain is underlined. (*) conserved residues among all EsPGRPs; (.) substitutions with weak similarity (<0.5 Gonnet PAM 250 matrix); (:) substitutions with a high similarity (>0.5 Gonnet PAM 250 matrix).

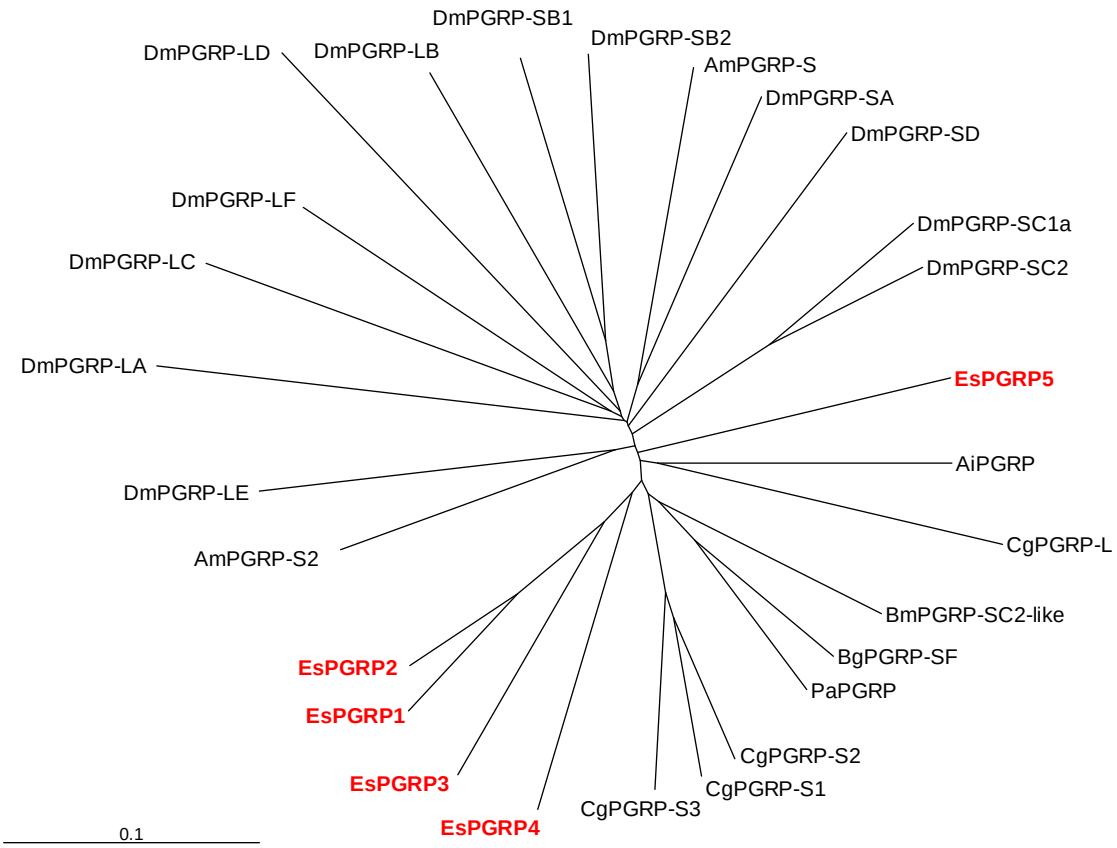

Figure 6. Neighbor-joining tree shows EsPGRP5 is the distant relative of previously described PGRPs from *E. scolopes.* Proteins were aligned with ClustalX using BLOSUM scoring matrix. Tree was constructed using neighbor-joining method as described in Troll et al. PGRPs shown from *E. scolopes* (Red), *D. melanogaster* (Dm), *A. melifera* (Am), *A. irradians* (Ai), *C. gigas* (Cg), *B. majavacas* (Bm), *B. globate* (Bg).

Table 4. Coverage of EsPGRP5 compared to other highly expressed genes.

Transcript	Putative Protein ID	e-value	Avg. coverage
Contig 1313	Beta actin	0	103.91
Contig 5159	Profilin	3e-72	28.50
Contig 1197	Translation elongation factor EF-1 alpha	0	17.66
Contig 1384	Cytochrome b	2e-42	15.01
Contig 1211	Heat shock protein 70	0	11.83
Contig 1364	EsPGRP5	5e-38	11.48
Contig 1222	Cytochrome c oxidase subunit I	3e-125	11.39
Contig 4119	Ribosomal protein 40S S2	9e-91	2.36

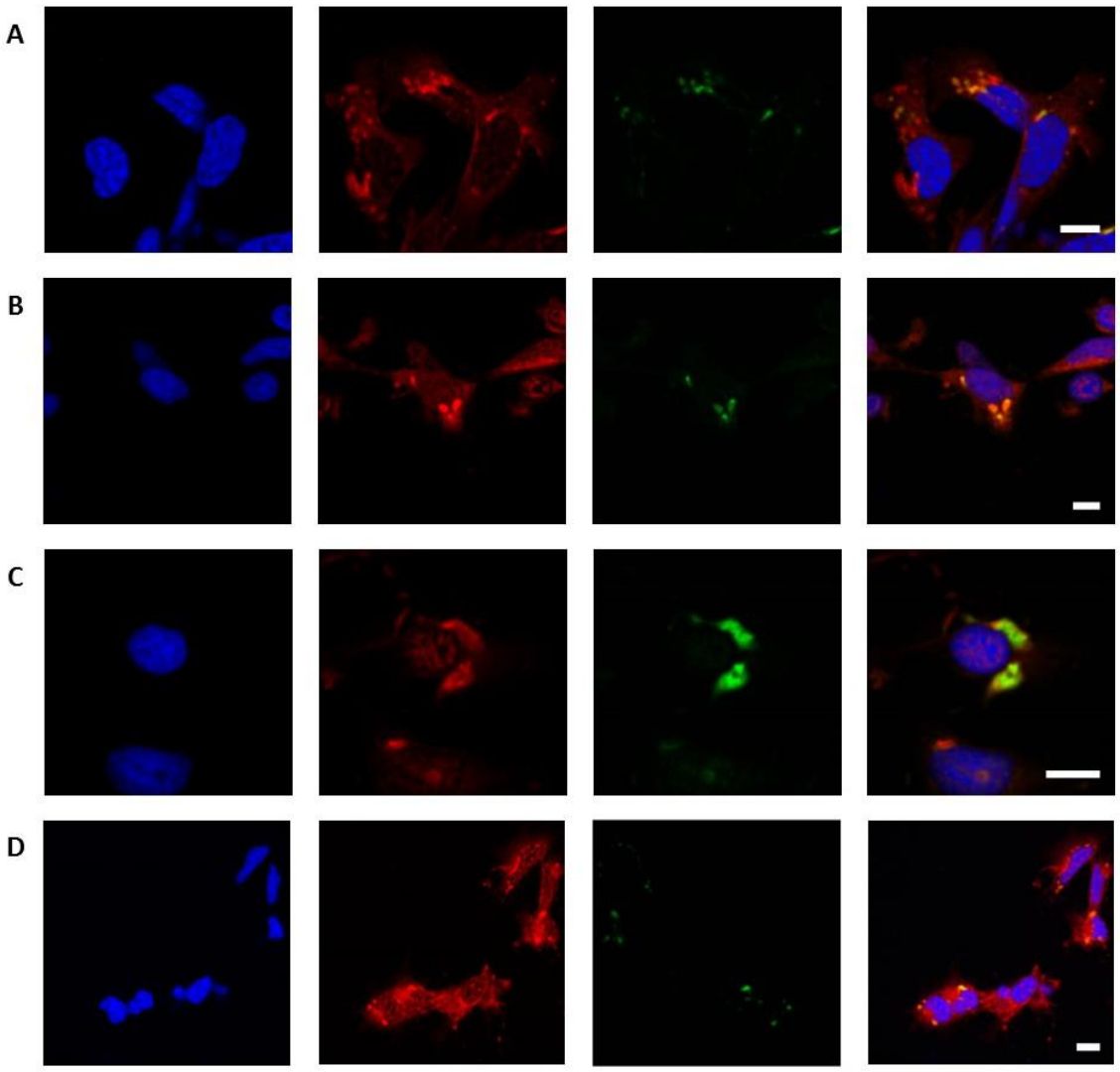

Figure 7. EsPGRP5 localizes to vacuoles within hemocytes of *E. scolopes.* (A) Hemocytes taken from symbiotic animals showed PGRP5 localized to vacuoles within hemocytes. These vacuoles seemed to colocalize with the rhodamine phalloidin staining, indicating possible colocalization with the actin cytoskeleton. This staining pattern did not change if the hemocytes were challenged with *V. fischeri* (B) or *V. harveyi* (C). Hemocytes taken from animals cured of their symbionts also had PGRP5 present in vacuoles within the cells (D). Scale bars indicate 10 µm.

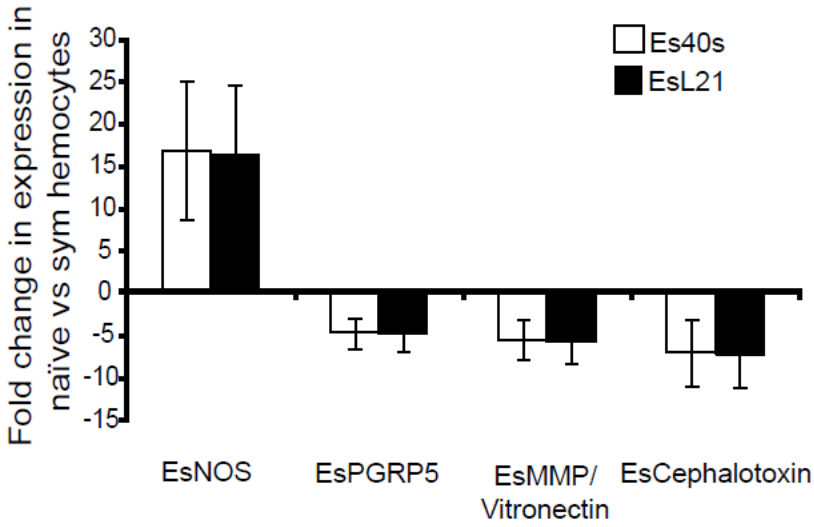

Figure 8. Differential expression of select immune genes in adult symbiotic and naïve hemocytes measured by qRT-PCR. Transcription of normalized genes was normalized to the 40S and L21 ribosomal subunit genes and represented as fold difference of the naïve relative to the symbiotic state. EsPGRP5 is down-regulated by nearly 5-fold in naïve hemocytes. Error bars represent standard deviation of three triplicate samples and is representative of three separate biological samples.

Chapter 3. The Accessory Nidamental Gland[2]

Introduction

Many aquatic and marine invertebrates, including some cephalopods (squid, octopuses, and cuttlefish), lay their eggs in clutches or masses on benthic substrates, where they take weeks or even months to develop before hatching (117-120). During this time, the developing embryos are unprotected and prior observations suggest these egg clutches resist predation and/or fouling by microorganisms, although clear mechanism for this resistance have yet to be described. Sexually mature females of some species have an accessory nidamental gland (ANG), a reproductive organ that houses a consortium of bacteria in pigmented epithelium-lined tubules and is attached to the nidamental gland (NG), the organ that secretes the jelly coat surrounding fertilized eggs (121). Culture-dependent and –independent methods have identified the dominant members of these microbial communities for some squid species (49, 121, 122). All squid ANGs examined to date are dominated by *Alphaproteobacteria*, usually members of the *Roseobacter* clade within the *Rhodobacterales* (48, 50, 121) with additional members belonging to the *Gammaproteobacteria* (vibrios, pseudoalteromonads, and pseudomonads) and the *Bacteroidetes*. Similar taxonomic groups were also found in the egg casings of the squid *Doryteuthis pealeii*, suggesting that the ANG serves to inoculate the egg clutches with a bacterial population (121). Although the exact role of these consortia has not been determined, those past studies suggest a symbiotic relationship between these bacteria and their hosts that should be investigated further.

[2] Collins AJ, LaBarre BA, Won BSW, Shah MV, Heng S, Choudhury MH, Haydar SA, Santiago J, Nyholm SV. 2012. Diversity and partitioning of bacterial populations within the accessory nidamental gland of the squid *Euprymna scolopes*. *Appl. Environ. Microbiol.* **78**:4200-8.

In this study, we examined the accessory nidamental gland of the Hawaiian Bobtail squid, *Euprymna scolopes* (Figure 1). The symbiosis between *E. scolopes* and the bioluminescent bacterium *Vibrio fischeri* is used as a model system to study the effects of beneficial bacteria on the development of animal host tissues (11, 21, 70, 123). Adult *E. scolopes* squid can easily be collected and bred in the laboratory and are readily accessible to use as experimental animals to research host/microbe interactions. In addition, its responses (e.g., biochemical, cellular, and developmental) to bacterial colonization are the best characterized for any cephalopod species.

In order to better understand the role of the ANG consortium in the biology of *E. scolopes*, the host and bacterial cell morphologies as well as the microbial diversity were characterized using transmission electron microscopy (TEM), 16S ribosomal sequence analysis, restriction fragment length polymorphism (RFLP) analysis, fluorescent *in situ* hybridization (FISH), and high-throughput 454 metagenomic sequencing. Here, we report the initial characterization of the ANG microbiota for the model host, *E. scolopes*. This study is the first to use high-throughput sequencing to characterize the bacteria in any accessory nidamental gland. More importantly, it sets the foundation for exploration of a bacterial consortium and a monospecific symbiosis within the same host.

Materials and Methods

Animal maintenance

Adult animals were collected from shallow sand flats of Oahu, HI, by dip net and maintained in 42-liter recirculating aquaria at the University of Connecticut with artificial seawater (ASW; Instant Ocean) at 23°C and kept on an approximately 12 h light / 12 h dark cycle (120). Adults were fed 3-4 ghost shrimp during the overnight hours.

Juvenile animals were raised in 3 gallon plastic livestock feed pans with autoclaved sand and 6 L of filtered ASW. Feedings were done with mysid shrimp were provided as a food source, adding 1-2 shrimp per animal for 30 minutes, then removing the uneaten shrimp (124). These feedings were repeated 4 times a day until animals were large enough to consume the ghost shrimp fed to the adult animals.

Dissection and DNA extraction

Female squid that had been maintained in the laboratory for between 24 h and 4 months were anesthetized in 2% ethanol in filter sterilized ASW and ventrally dissected to remove the accessory nidamental gland. Once removed, the ANGs were flash frozen in liquid nitrogen and stored at -80°C until use. For ANG clone libraries, ANGs were homogenized in lysis buffer and total DNA was isolated using a DNeasy Tissue Prep kit (Qiagen, Hilden, Germany). To obtain DNA for 454 library construction, frozen ANGs were first thawed and then homogenized in Squid Ringer's solution (530 mM NaCl, 25 mM MgCl2, 10 mM CaCl2, 20 mM HEPES, pH = 7.5) using a ground glass homogenizer. The homogenate was spun for 10 min at 5,000 x g at 4°C. To remove solubilized host tissues, the supernatant was removed and the pellet was repeatedly washed (at least three times) with squid Ringer's solution until the protein concentration of the supernatant was sufficiently low (< 0.5 mg/mL), as measured spectrophotometrically by A_{280} analysis. For 454 sequencing, total genomic DNA was extracted from the resulting pellet by the use of a DNA MasterPure kit (Epicentre, Madison, WI).

Culturing bacteria from the ANG

Frozen ANGs from three sexually mature squid were homogenized in squid Ringer's solution, and the homogenate was serially diluted 10-fold and plated in triplicate onto Reasoner's 2A media (125) supplemented with 27 g of marine salts (Instant Ocean) or salt water tryptone (SWT). Plates were incubated aerobically at 28°C for 3 days and the resulting colonies were observed for pigmentation.

Microscopy

Immediately after dissection, ANGs were cut in half and fixed at room temperature in 2.0% paraformaldehyde-2.5% glutaraldehyde in buffer A (0.1 M sodium cacodylate, 0.375 M NaCl, 1.5 mM CaCl2, and 1.5 mM MgCl2, pH = 7.4). After an initial 15-min fixation, the tissue samples were cut into smaller pieces (~0.25 cm thick) and placed in fresh fixative for an additional 5 h at 4°C. Following fixation, tissue pieces were washed several times in cold buffer A and left at 4°C overnight. The following day, tissues were post-fixed in a solution of 1% osmium tetroxide-0.8% potassium fierricyanide-0.1 M sodium cacodylate-0.375 M NaCl for 1.5 h at 4°C and then washed in distilled water, dehydrated through an ascending ethanol series, cleared in 10% acetone and embedded in an epoxy mixture of Embed 812 (Electron Microscopy Sciences, Hatfield, PA) and Araldite 506 (Ernest Fulhan Inc., Albany, NY). Semithin (2 µm) sections were cut with a glass knife using an LKB Ultramicrotome V and stained with methylene blue and azure II followed by counterstaining with basic fuchsin. Stained sections were viewed on an Axiovert 200M (Zeiss, Oberkochen, Germany) microscope. Thin (80 nm) sections were cut using a diamond knife on a LKB Ultramicrotome V followed by staining with 2% uranyl acetate and Reynolds's lead citrate (126) and

viewed with an FEI Tecnai Biotwin G2 Spirit electron microscope (Hillsboro, OR) operated at 80 kV.

16S clone library construction and RFLP analysis

To examine the bacterial diversity in ANGs, total genomic DNA from the organs of five sexually mature females were used to make five separate 16S clone libraries. ANG1, ANG2, ANG3, and ANG4 came from four females that were kept in our squid facility for 9, 14, 12 and 17 weeks, respectively. ANG5 came from an individual that had been field caught and was maintained in our facility for 24h. 16S genes were amplified using 25 µl GoTaq (Promega, Madison, WI) PCR reactions with 27F and 1406R eubacterial 16S primers (Table 1). PCR conditions were as follows: 95°C for 3 min, then 35 cycles of 95°C for 30 s, 57°C for 30 s and 72°C for 90 s followed by a final elongation for 72°C for 10 min. PCR products were ligated and cloned using a PGEM-T Easy kit with JM109 cells (Promega, Madison, WI). A total of 417 colonies were selected for restriction fragment length polymorphism analysis by incubating cloned genes with 10 U of MspI restriction enzyme (New England Biolabs, Ipswich, MA) at 37°C for 15 min. The resulting fragments were visualized on a 1.5% agarose gels and clones were grouped according to unique RFLP patterns. Representative clones from each group were sequenced using BigDye version 1.1 (Applied Biosystems, Carlsbad, CA) according to the manufacturer's specification. Any clone that could not be assigned to an RFLP group was also sequenced. The 16S rRNA genes from 25, 45 and 27 clones from ANG1, ANG2 and ANG3, respectively were fully sequenced to confirm the accuracy of the restriction digest grouping. Sequences were analyzed with the Bellerophon chimera server (127) and 25 chimeric sequences were discarded, leaving 392 clones that were

included in the analysis. The full-length sequences were used to search the Greengenes 16S rRNA gene database of named isolates by the use of BLAST (128). Operational taxonomic units (OTUs) were assigned to each sequence based on the highest percent identity. Sequences from the *Verrucomicrobia* had few quality and alignments and were therefore characterized at the phylum level.

454 Metagenomic Sequencing

To identify other bacterial members from the ANG that might not have been detected with 16S clone libraries and to increase our sequencing depth, we analyzed bacterial diversity using metagenomic analyses. Bacterial DNA was extracted from 3 ANGs as described above. The samples were pooled, and 500ng was used to construct a 454-shotgun metagenomic library using a Rapid Library kit (Roche Applied Science, Basel, Switzerland). After the small-volume (SV) emulsion PCR (emPCR) was performed, the library was used in two 454 sequencing runs with FLX Titanium chemistry. After removing artificial replicates by the use of a 454 replicate filter (129), 622, 987 sequences with an average length of 389.68 bases (a total of 242.77mB) were analyzed. Roughly 1% of the reads (6,350) were eukaryotic in origin and not used for our analysis.

For 16S analysis of 454 data, reads were annotated using the MG-RAST server (130). Using the algorithm available from the Ribosomal Database Project (RDP), reads with at least a 200-bp alignment to a known 16S gene were extracted and used to search the Greengenes 16S database with BLAST. OTUs were assigned as described above.

16S rRNA FISH

To localize bacteria within the ANG, organs were dissected from six sexually mature female squid and prepared for fluorescent in situ hybridization (FISH). Two were freshly collected and dissected in Hawaii; the other four were kept in our animal facility for 8 to 14 weeks prior to dissection. Time in captivity did not affect results (data not shown). Three ANGs were fixed with Carnoy's solution (ethanol:chloroform:acetic acid [6:3:1]) overnight and the other three were fixed in 1X PBS-4% paraformaldehyde for 4 h. Tissues were embedded in paraffin and hybridization was performed as previously described (15). Three egg capsules were removed from freshly laid egg clutches, fixed in squid Ringer's solution-4% paraformaldehyde for 4 h and embedded in paraffin as described above.

Based on our 16S data, several ribosomal probers were used at 50 pmol/mL each for hybridization (Table 1). Probes for species of Bacteria, *Alphaproteobacteria*, the *Roseobacter* clade or *Cytophaga-Flavobacteriia-Bacteroidetes* (CFB) were taken from previous studies. Novel 16S probes for *Verrucomicrobia* and *Phaeobacter* species were designed based on 16S sequence data and specificity was confirmed with ProbeCheck (131) and fixed cultures of closely related members of genera of *Rhodobacterales* (e.g. *Phaeobacter*, *Ruegeria*, *Tateyamaria*, and *Nautella* for the *Phaeobacter*-specific probe; data not shown). All probes were synthesized by Eurofins MWG Operon (Huntsville, AL) and conjugated to fluorescein isothiocyanate (FITC), Cy3 or Cy5. After an overnight hybridization at room temperature in formamide hybridization buffer (0.9 M NaCl, 20 mM TRIS [pH 8], 0.01% sodium dodecyl sulfate [SDS]), the tissue was washed in hybridization buffer and then counterstained with 300 nM of DAPI (4', 6-diamidino-2-

phenylindole) (Invitrogen, Carlsbad, CA) in 1X PBS for 5 min. The following negative controls were performed: no probe, a nonsense probe (complimentary to the eubacterial probe Eub338) and competition with unlabeled probes. Tissue sections were imaged on a Leica SP2 confocal microscope (Wetzlar, Germany) or a Zeiss Axiovert 200M epifluorescence microscope (Carl Zeiss, Germany) using DAPI, FITC, Cy3 and Cy5 filter sets.

Enumeration of Bacteria by DAPI Counting

In order to estimate the total bacteria within the ANG, organs from three sexually mature females were dissected and homogenized as described above. To separate the bacterial cells from host tissue, the homogenate was spun at 5,000 x g for 10 min. The supernatant was removed and the pellet was resuspended in squid Ringer's solution. This was spun at 500 x g for 10 min. The supernatant was removed and centrifuged at 5,000 x g for 3 min. After discarding the supernatant, the pellet was resuspended in 1 mL of 4% paraformaldehyde in squid Ringer's solution and incubated overnight at 4°C. The cells were diluted 1,000 or 10,000-fold in water containing DAPI (20 µg/mL) and 0.01% Triton and incubated for 5 minutes at room temperature before being applied to a black 0.2 micron filter. Filters were viewed with a fluorescent microscope and 20 fields of view were counted and the number of bacteria in the original homogenate was determined using the following formula: (Dilution factor x average cells per field x area of filter) / (area of field x volume filtered).

Accession numbers

16S clone library sequences were deposited in the European Nucleotide Archive (ENA) with accession numbers HE574851 to HE574928. Metagenomic reads were

deposited in the NCBI Short Read Archive (SRA) with accession numbers SRR329677.8 and SRR329678.5.

Results

Morphological and EM observations

The ANG of *E. scolopes* (Figure 1a) contains many convoluted and pigmented tubules (Figure 1b). While most tubules have a dark orange pigmentation, some appear white (Figure 1c) or, more rarely, yellow or red (not shown). As with other ANGs, the bacteria within the tubules likely synthesize these pigments as colonies isolated from the organ also appeared similarly pigmented when grown in culture (Figure 1c, 121)

Light and electron microscopy of fixed sections of the ANG revealed that the organ is highly vascularized, with many blood vessels among tubules lined with ciliated epithelial cells and containing bacteria. In some tubules, however, bacteria were not observed (Figure 2A). Two morphologically distinct and segregated cell types were observed (Figure 2B): a large coccoid bacterium (LCB) and a smaller bacillus bacterium (SBB). These two bacterial morphotypes appeared in separate tubules with strikingly different epithelia. One type of epithelium, associated with the SBB, appeared vacuole-rich (Figure 2B and C), while the other, associated with the LCB, had an electron-dense staining pattern lacking vacuoles (Figure 2B and F). Within the tubules housing the bacteria were microvillar brush borders 1 to 5 µm in thickness along with membrane-bound vesicles that may be secreted or blebbed from the host (Figure 2D). Some of the vacuole-rich epithelial cells had the distinct appearance of being secretory in nature, containing numerous large electron-light vacuoles and smaller electron-dense granules located at the apical surfaces of the epithelium (Figure 2C). Hemocytes, the primary

innate immune cells of *E. scolopes*, were observed in the lumina of the tubules; however, phagocytosed bacteria were not observed within these cells (Figure 2E). Each tubule was dominated by one of the two morphologies: either the SBB or the LCB.

Mixtures of both LCB and SBB morphotypes were also observed outside the tubules within the connective tissue (Figure 2G and 2H). The epithelial membranes appeared well-preserved, suggesting that these observations were not from a fixation artifact and that the bacteria can travel outside the ANG lumina. Hemocytes were also observed within the connective tissue (Figure 2H), but as in the lumina of the tubules, no intracellular or phagocytosed bacteria were noted. Under higher magnification, the LCB cells appeared to be filled with many granules (Figure 2I). This was in stark contrast to the SBB which were either mostly electron dense (Figure 2J) or contained large electron-light storage vacuoles which resembled polyhydroxybutyrate (132, 133, Figure 2K).

Bacterial Load in ANG

Enumeration of bacteria by DAPI counts showed that bacteria within this organ can range roughly from 5×10^9 to 5×10^{10} cells per organ (Figure 3). Of the three animals counted the lowest count seen was 6.2×10^9 cells per organ and the highest was 6.2×10^{10} cells per organ. Plate counts of culturable isolates from the ANG also tend to be in this range, suggesting that 2×10^9 culturable cells exist in an ANG (not shown).

16S Diversity

In order to identify members of the microbial community within the ANG of *E. scolopes*, we constructed five 16S clone libraries from five sexually mature adult female squid. Sequences of 417 clones were binned by RFLP analysis. Of these, 96 full-length 16S sequences were analyzed and 25 identified as chimeric sequences were removed.

Analysis of these data showed that most (302/392); Table 2) clones belonged to the *Alphaproteobacteria* and that *Phaeobacter* was the most commonly observed genus (221/392). Other *Alphaproteobacteria* species belonged to genera within the *Rhodobacterales*, primarily of the *Roseobacter* clade (e.g. *Ruegeria, Silicibacter, Phaeobacter*, etc.). Twelve were from clones were from the genus *Kordiimonas*, and eight sequences were from the *Rhizobiales*. The next most common group of sequences had greatest similarity to members of the phylum *Verrucomicrobia*. However, these sequences displayed very low identity to characterized isolates (~90% identity). This is most likely due to a lack of characterized marine isolates from the phylum *Verrucomicrobia*. Only one sequence belonged to the *Gammaproteobacteria* and shared greatest identity to *Shewanella* 16S sequences.

The overall bacterial populations of ANGs from separate animals were similar with two OTUs conserved across all five clone libraries, *Phaeobacter* and *Verrucomicrobia*. Members of four other genera of the *Alphaproteobacteria* (*Ruegeria, Kordiimonas, Cohaesibacter* and *Nautella*) were conserved among the same four ANG libraries. Length of time spend in the mariculture facility did not seem to influence the microbial communities found in the ANG, as animals maintained for either 1 day (library ANG5) or 4 months (library ANG4) were found to have similar bacterial taxa.

In addition to the RFLP and sequence data from the five clone libraries, 16S gene fragments from the 454 metagenome were also analyzed. A total of 532 genomic fragments with at least a 200bp alignment to a reference 16S sequence in the RDP database were used for this analysis. The taxonomies of these 16S sequences were similar to those identified in the 16S libraries; species of *Rhodobacterales, Rhizobiales* and

Verrucomicrobia were dominant (Figure 4). 72.55% of the 16S sequences belonged to the *Alphaproteobacteria*, and the most common genus was *Phaeobacter* (177/532). Members of the order *Rhodobacterales* was the most common, with *Verrucomicrobia* being the second largest contingent overall (89/532). Members of the *Rhizobiales* and the phylum *Bacteroidetes*, which were not seen our 16 libraries, accounted for less than 20% of the 16S sequences. Only 2 of the 532 16S sequences were from the *Gammaproteobacteria*.

FISH

Observations from the electron microscopy suggested that different morphotypes (SBB or LCB) dominated individual tubules within the ANG (Figure 2). To test whether this could have been due to different phylogenetic groups occupying separate tubules, fluorescent in situ hybridization (FISH) was used to visualize dominant bacterial taxa within the ANG. Ribosomal FISH revealed that most tubules within the ANG contained a specific bacterial group (Figure 5). Staining with a FITC-labeled *Roseobacter* clade specific probe and a cocktail of the eubacterial probes Eub338I and Eub338III showed that the *Roseobacter* clade probe hybridized to the majority of the bacteria of one tubule (Figure 5a). Similarly, using both the *Verrucomicrobia*- and *Alphaproteobacteria*-specific probes, tubules were dominated by only one of the two fluorescent signals, suggesting bacterial partitioning among the ANG tubules (Figure 5b). The presence of members of members of the *Cytophaga-Flavobacteriia-Bacteroidetes* (CFB) that were identified in the 454 metagenomic sequencing analysis was also confirmed with this technique (Figure 5c).

Applying FISH to the jelly capsule of freshly laid squid eggs also revealed a mixture of bacteria within the capsule, with an abundance of *Alphaproteobacteria* present, including species of *Phaeobacter* (Figure 4d and e). No bacterial cells were observed in direct contact with the developing embryo (not shown). These data suggest that the bacteria from the ANG are deposited directly into the host egg capsules.

Juvenile Squid ANG development

Three cohorts of juveniles were successfully raised to maturity (mantle length > 20 mm). Ten of these animals were females and dissected to remove their ANGs. Surprisingly, 7 of the animals had no ANG development (Fig. 6A&B). Only 3 of the females raised in the artificial seawater system contained an ANG and these were substantially underdeveloped (Fig. 6C). Organs were less than 1/3 the shape of a mature ANG and were unsymmetrical. Of these 10 females, one female laid a single clutch, which was small and the embryos in this clutch failed to develop.

Discussion

We used a variety of microscopy and molecular methods to characterize the bacterial populations of the ANG of *E. scolopes*, ultimately to understand its role in host reproduction. These analyses show that *Alphaproteobacteria* species from the *Roseobacter* clade with the *Rhodobacterales* are prevalent and that members of the genus *Phaeobacter* dominate the consortium, while other major constituents are members of the *Rhizobiales*, *Verrucomicrobia*, and *Flavobacteriia*. This bacterial consortium is contained within heteromorphic epithelium-lined tubules that are infiltrated by host hemocytes. Moreover, FISH analyses confirmed that many tubules of the ANG are dominated by single taxonomic groups, suggesting niche specificity in this association.

The dominance of the *Roseobacter* clade members in the ANG of *Euprymna scolopes* is similar to what has been described for other cephalopod ANGs, including those of other squid (49, 51) and cuttlefish (48), by the use of 16S clone sequencing. Like that of *Doryteuthis pealeii*, the ANG of *Euprymna scolopes* has a large *Alphaproteobacteria* contingent, comprising *Roseobacter* clade members as well as members of the marine *Rhizobiales*. The *E. scolopes* ANG also has a *Flavobacteriia* contingent, similar to observations made using egg casings of *D. pealeii* (49).

The presence of *Verrucomicrobia* and the lack of *Gammaproteobacteria* make the consortium in *E. scolopes* strikingly different from the ANG consortia previously described for other cephalopods. Members of the *Verrucomicrobia* have been detected in relatively few host/microbe associations (134-137) and the major presence of this group in the squid ANG represents a potentially novel symbiotic role for this phylum.

Less than 1% of 16S genes from our clone libraries and the 454 metagenome belonged to the *Gammaproteobacteria*. This is surprising for a number of reasons. In *D. pealeii*, it has been estimated that 5% of the bacterial population of the ANG is made up this group (49). Furthermore, *E. scolopes* has a binary association with the bioluminescent gammaproteobacterium, *Vibrio fischeri* (11, 21, 70). Given that the host expels 10^6 to 10^7 symbionts from its light organ as part of a daily rhythm (23, 113), the close proximity of the two organs, and that the ANG consortium is likely environmentally transmitted, it is surprising that *V. fischeri* was not detected in our analyses.

Previous work has shown that the bacterial consortia within cephalopod ANGs are likely established by horizontal/environmental transmission (138). In that work,

Kaufman *et al.* examined development of the ANG in *Doryteuthis opalescens* and found that the organ develops 11 weeks after hatching and that colonization is likely due to horizontal/environmental transmission. This conclusion is also supported by the observation that the nearest relatives of the ANG isolates from *D. pealeii* are environmental strains (49). The ANG of *E. scolopes* is absent at hatching, and females tend to reach sexual maturity within 60 days (139). Therefore, horizontal transmission of the symbionts of the *E. scolopes* ANG is also probable. Field-caught animals at different stages of development of ANG symbiosis will be used for future analyses of both the organ and the microbial community.

Animals raised in an artificial seawater system showed underdeveloped or no developed ANG. Moreover, these animals had a dramatic decrease in fecundity. Although the females were observed mating with male squid, the females failed to produce any viable offspring. The absence of normal environmental bacteria may play a key role in the development of the ANG. Kaufman *et al*. showed that the ANG in *D. opalescens* acquires bacteria from the surrounding environment by sequestering bacteria within invaginations of the primordial ANG tissue which eventually completely encapsulate the bacteria and mature into fully-formed tubules (138). Given that these squid were raised in water without a natural bacterial population, we hypothesize that the presence of the ANG symbionts in the natural seawater play a role and may even induce the maturation of the ANG.

The data presented here suggest that establishment and maintenance of the bacterial consortium may be an intricate process, as both electron microscopy and FISH analyses showed bacterial partitioning among the ANG tubules. Electron microscopy

revealed distinct morphotypes (LCB and SBB) prevalent in each tubule, and FISH revealed that members of the *Rhodobacterales, Cytophaga-Flavobacteriia-Bacteroidetes,* and *Verrucomicrobia* dominated separate tubules. Other studies have noted dominant morphotypes (coccus and bacillus) within the lumina of other cephalopod ANGs. Electron micrographs from the cuttlefish *Sepia officinalis* show a coccoid bacterium with a morphology very similar to that of the granular, coccoid cells observed in the *E. scolopes* ANG (140). The data from this study suggest that the different bacterial morphotypes are different taxa occupying separate tubules. Bacterial morphology by itself is not a reliable taxonomic identifier; however, studies of *Coraliomargarita akajimensis,* a marine *Verrucomicrobia* species, revealed a morphology very similar to the LCB morphotype observed in this study (141). The two epithelial morphologies of the tubules are very distinct from one another, suggesting that each tubule fosters a unique microenvironment optimized to contain a specific bacterial taxon or that specific bacteria influence development of different epithelia. The mechanism(s) for establishing and maintaining bacterial tubule dominance is not yet known, but these bacterial groups may be adapted to specific niches or microenvironments within the ANG. Alternatively, different taxa may dominate specific tubules during colonization due to a founder effect.

Just as carbon and energy sources influence bacterial diversity in digestive tracts, nutrition may play a role in the segregation of the bacteria within the ANG. While members of the *Verrucomicrobia* have not been thoroughly described, many have been shown to degrade polysaccharides such as mucin in the human gut (142) or fucoidan in the gut of a sea cucumber (143). The genome of *Phaeobacter sp.* ANG1, a dominant member of the ANG consortium in *E. scolopes,* reveals that it has many pathways for

energy and carbon assimilation; however, it lacks enzymes to degrade polysaccharides, including chitinases, amylases, agarases, and α-L-fucoidase (144). Therefore, the host may provide different nutrients in different tubules, thereby enriching for certain bacteria. The presence of PHB-like granules in some bacterial cells could be explained by nutrient restriction, as PHB improves survivability during starvation and/or stress tolerance in other systems (133, 145).

There are other clues that can be gathered from the host as to how a microenvironment can be gathered from the host as to how a microenvironment can be created to foster dominance of specific bacteria. The light organ symbiosis between *E. scolopes* and *V. fischeri* has been studied in detail for more than 20 years, and previous studies have shown that the host and symbiont work in concert to create a microenvironment that selects for *V. fischeri* to the exclusion of nonsymbiotic bacteria (11, 21, 24). The stark differences in epithelial tissues in the ANG suggest that unique microenvironments exist in the lumina of the tubules. In the light organ, hemocytes, representing the sole cellular component of the host's innate immune system, have been implicated in establishing and maintaining specificity (23, 33, 34, 56). Hemocytes were also observed to infiltrate the lumina of the ANG tubules and were found to come in direct contact with the bacterial consortium. Whether these hemocytes contribute to specificity in the ANG association remains to be determined, but we have isolated several ANG bacterial strains that are available to use in adhesion and phagocytosis assays. Future research should examine how components of the innate immune system as well as other host and symbiont factors influence the development and maintenance of this association.

Despite numerous studies that have characterized the bacterial communities within cephalopod ANGs, its function is still unknown. The ANG or its bacterial symbionts may provide antimicrobial or antifouling compounds that protect the squid's eggs throughout their development (146, 147). The genome of *Phaeobacter* sp. ANG1 revealed no classical antibiotic synthesis pathways (144) including antimicrobials associated with the *Roseobacter* clade, such as tropodithietic acid and indigoidine (44, 148). Future analyses should look for antimicrobial activity in ANG isolates, as novel biosynthetic pathways of antimicrobials will not be found with bioinformatic methods. Future studies should test whether removing the ANG consortium from female squid and/or their egg clutches influences fecundity and egg development. Characterizing more isolates from the ANG, particularly the symbionts from the *Bacteroidetes* and *Verrucomicrobia* should also shed light on the function of the ANG.

The female members of many squid species found worldwide harbor a consortium of bacteria within their ANG. Surprisingly, much of the composition of the microbial communities is the same (e.g. the dominance of *Rhodobacterales*), even though these hosts are found in very different environments with different physical and biological parameters (e.g. salinity, temperature, predators, and life histories). This trend suggests that these microbial consortia play similar roles in their squid hosts. Future experiments can utilize high-throughput sequencing techniques to reveal gene expression of the bacteria in the ANG and within the egg capsule. The results of this study lay the foundation for the development of *E. scolopes* as a model for studying consortial symbiosis (ANG) and a binary symbiosis (light organ) in the same host.

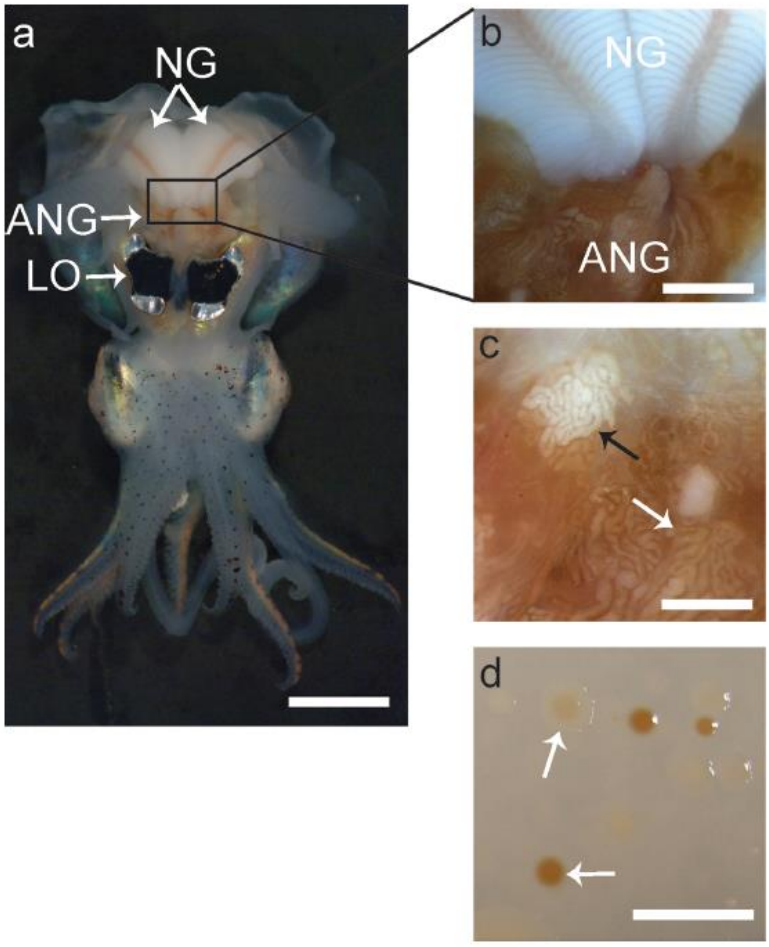

Figure 1. Anatomy of a female *Euprymna scolopes* and morphology of ANG isolates. (a) Ventral dissection of E. scolopes, showing the accessory nidamental gland (ANG) located posterior to the light organ (LO) and in close proximity to the nidamental gland (NG). (b) Pigmented ducts in the NG converge at the ANG. (c) Magnification of the ANG reveals convoluted tubules, most of which are dark orange in pigmentation (white arrow), but others appear white (black arrow) or yellow (not shown). (d) Culturing of ANG symbionts results in many colonies with pigmentation similar to that of the ANG tubules (white arrows). Bars, 1cm for panes a and d, 2.5 mm for panel b and 1 mm for panel c.

Table 1. Primers and FISH probes used in this study

Primer, probe or target organisms	Name	Sequence	Hybridization buffer % formamide (probes only)	Source
Eubacterial	27F	AGAGTTTGATCCTGGCTCAG		149
Eubacterial	1406R	ACGGGCGGTGTGTRCAA		149
FISH Probes				
Eubacterial	Eub338I	GCTGCCTCCCGTAGGAGT	30	150
	Eub338III	GCTGCCACCCGTAGGTGT	30	151
Roseobacter	G Rb	GTCAGTATCGAGCCAGTGAG	30	152
Bacteroidetes	CF319	TGGTCCGTGTCTCAGTAT	30	153
Verrucomicrobia	Verruco_193	CGCCATTACAAGCTTTAGTA	20	This study
Phaeobacter	Phaeo_126	TGGCTATTTTAGAGAAGGGCA	20	This study
Alphaproteobacteria	Alph_968	GGTAAGGTTCTGCGCGTT	30	154
Eubacteria (neg. control)	NonEub338	ACTCCTACGGGAGGCAGC	30	155

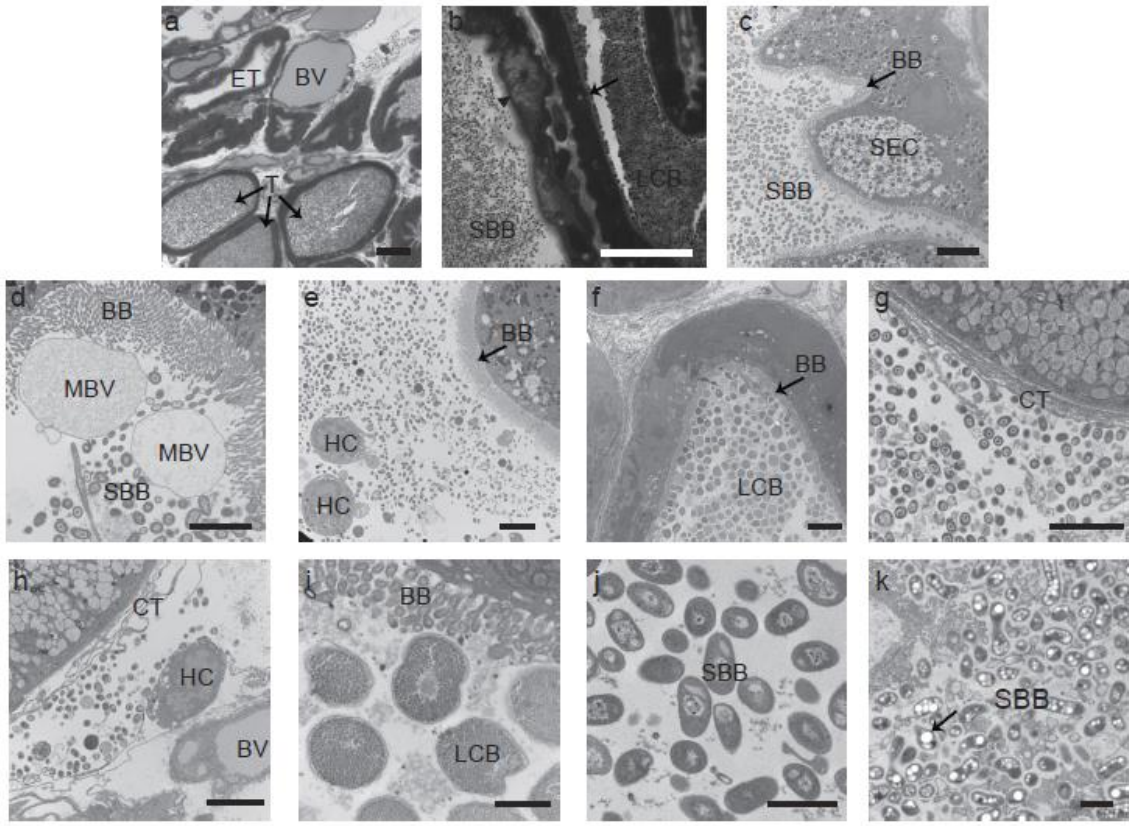

Figure 2. Light microscopy and TEM of fixed sections from the *E. scolopes* ANG. (a) Cross-section of ANG tissue, showing many tubules interspersed between blood vessels (BV). Most tubules (T) contained dense bacterial populations; however, some were empty (ET). (b) Closer inspection revealed that tubules had two different bacterial morphotypes, a large coccoid bacterium (LCB) and a smaller bacillus-shaped bacterium (SBB) that were observed in tubules with differing epithelial morphologies, either a vacuole-rich epithelium (arrowhead) or dense epithelium (black arrow). (c) Tubule with brush border (BB) dominated by SBB with an epithelial morphology suggesting secretory cells (SEC). (d) SBB inhabiting tubules with a thick microvillar brush border (BB) and membrane bound vesicles (MBV). (e) Hemocytes (HC) in the lumen of the tubules with a population of bacteria. (f) Tubule dominated by LCB. (g) Bacteria were also seen in the connective tissue (CT) outside the tubules: note the absence of a brush boarder. (h) Hemocytes were also seen with bacteria among the connective tissue. (i) Closer inspection of the LCB revealed many storage granules. (j and k) One SBB morphotype showed a nucleotide structure, while the other (k) showed many polyhydroxybutyrate-like granules (black arrow). Bars, 30 µm (a and b), 5 µm (c to h) and 1µm (i to k).

Table 2. Operational taxonomic units within five ANG 16S clone libraries

	No. of OTUs				
Phylotype	ANG1	ANG2	ANG3	ANG4	ANG5
Alphaproteobacteria					
Rhodobacterales					
Phaeobacter	20	49	78	40	26
Ruegeria	9	6	0	6	4
Nautella	2	4	0	5	1
Labrenzia	0	6	0	1	0
Pseudoruegeria	0	1	0	0	0
Oceanicola	0	1	0	0	0
Marinovum	0	0	2	0	0
Salipiger	0	0	4	0	0
Rhizobiales					
Cohaesibacter	0	1	0	2	3
Mesorhizobium	0	0	2	0	0
Kordiimonadales					
Kordiimonas	1	4	0	3	4
Verrucomicrobia	37	12	3	25	29
Gammaproteobacteria					
Shewanella	0	1	0	0	0
Total no. of clones	69	85	89	82	67

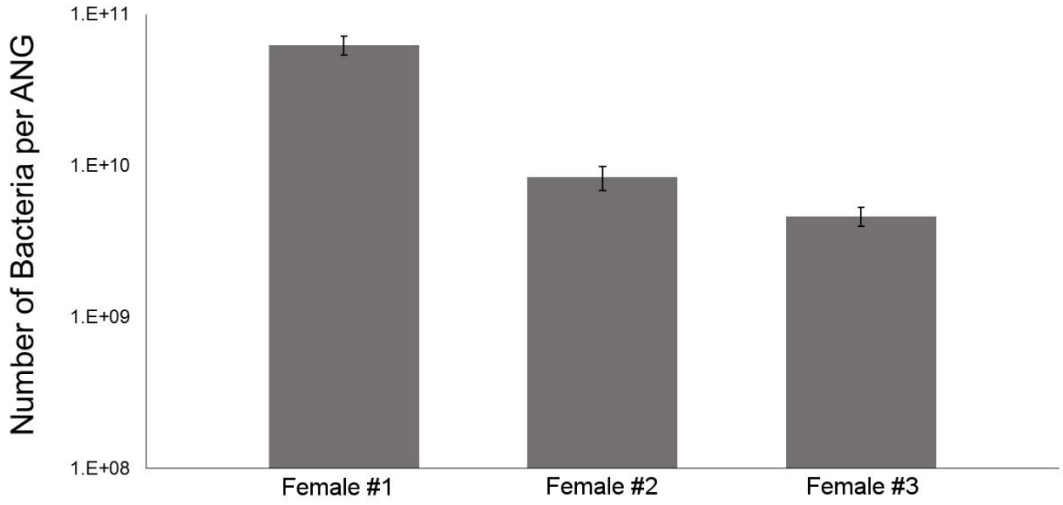

Figure 3. Bacterial counts by direct observation from three mature female ANGs. Enumeration of bacteria in the accessory nidamental gland suggests 5 x 10^9 to 5 x 10^{10} bacteria could inhabit a single ANG.

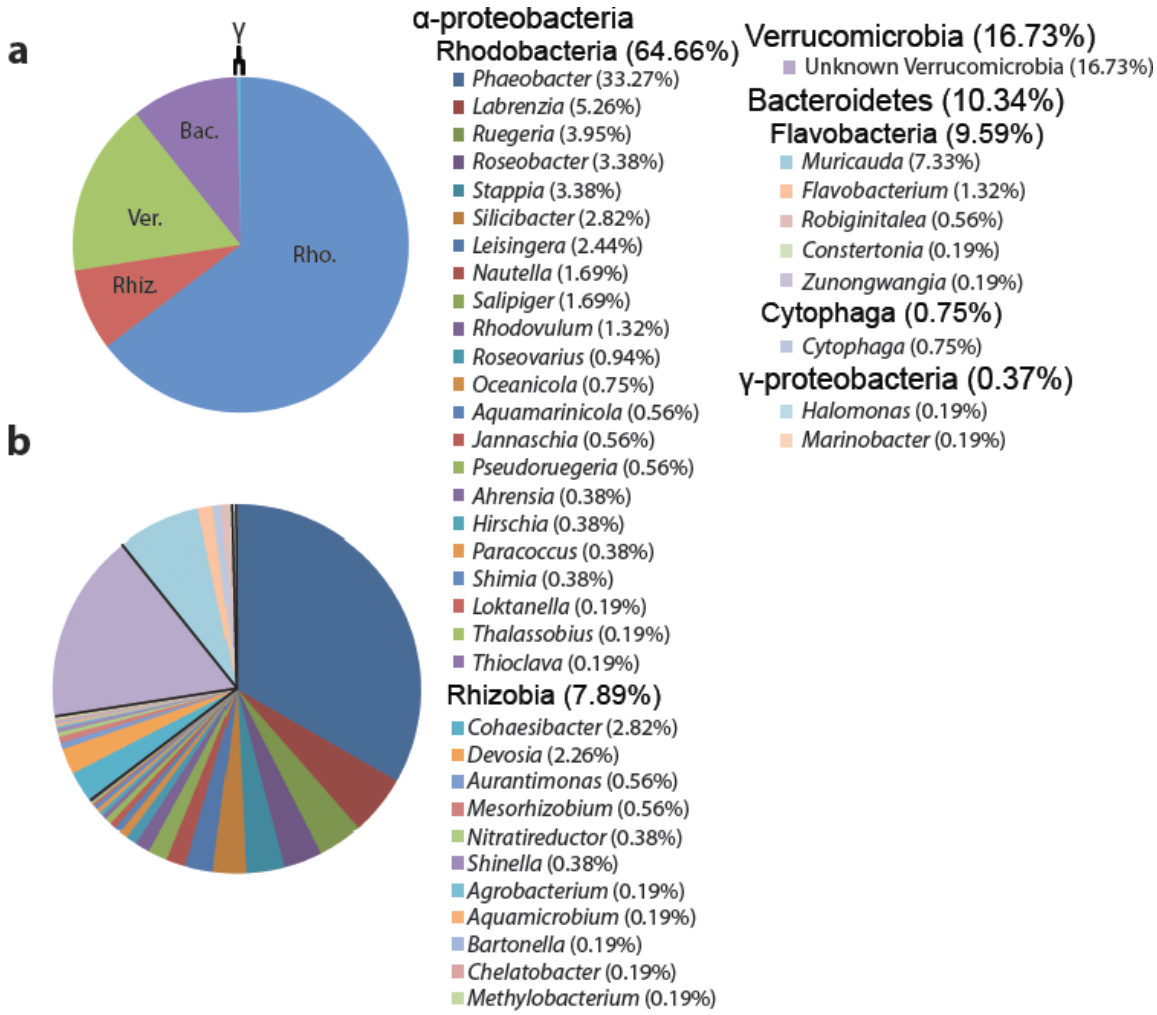

Figure 4. Taxonomic analysis of 454 metagenomic data. Higher (a) and lower (b) taxonomy of 532 16S gene fragments found within the ANG metagenome. Members of the *Alphaproteobacteria*, comprising mostly *Rhodobacterales* (Rho.), accounted for the vast majority of sequences (72.55%). Members of the *Verrucomicrobia* (Ver.) were the next largest contingent (16.73%), and members of the *Bacteroidetes* (Bac.) mostly from the *Flavobacteriia*, made up a smaller fraction (10.34%). Less than a 1% of 16S fragments belonged to the *Gammaproteobacteria*. The presence of a single 16S sequence is represented by "0.19%". Rhiz., *Rhizobiales*.

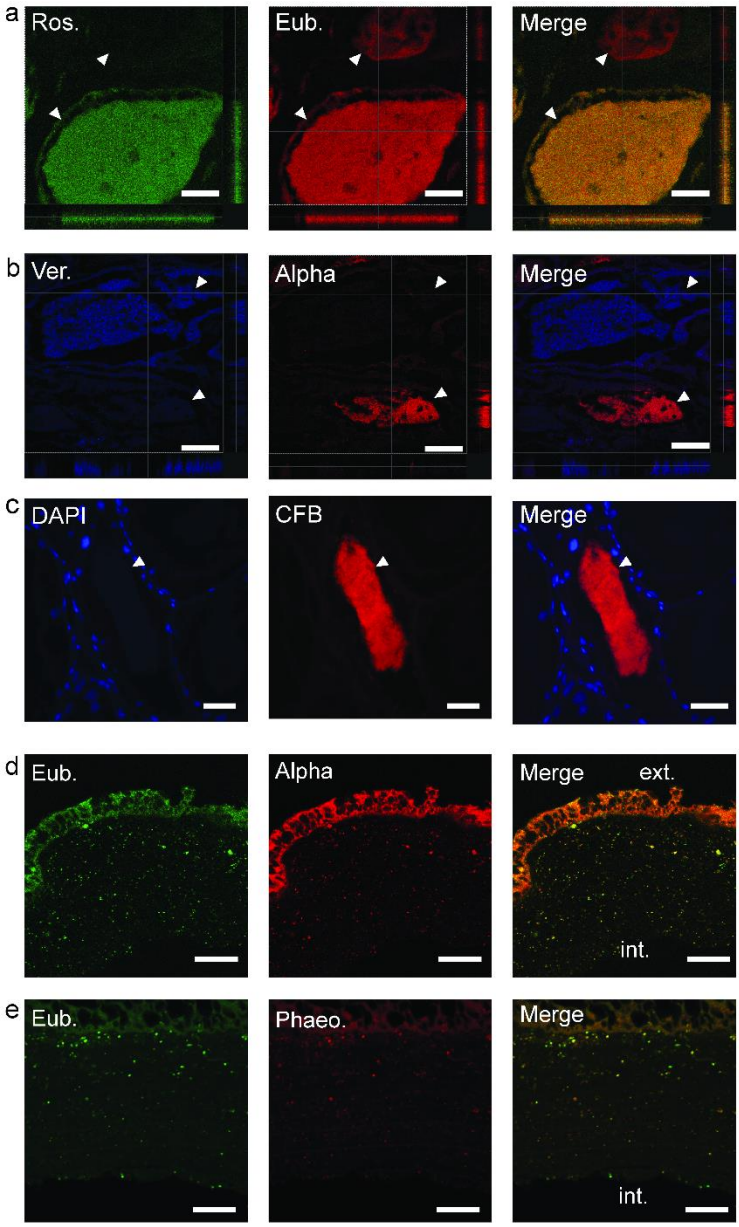

Figure 5. Fluorescent *in situ* hybridization of fixed ANG paraffin-embedded sections. (a) 16S FISH with FITC-conjugated *Roseobacter* (Ros.) probe (green) and CY3-conjugated eubacterial (Eub.) cocktail (red). (b) Cy3-conjugated *Alphaproteobacteria* (Alpha: red) and Cy5-conjugated *Verrucomicrobia* (Ver.; blue) probes were observed dominating separate tubules, suggesting specificity and/or segregation of the bacterial populations. Axes denote positions of "slices" through confocal sectioning. (c) 16S FISH with Cy3-conjugated *Cytophaga-Flavobacteriia-Bacteroidetes* probe (CFP; red) and DAPI staining (blue). (d) Hybridization with a FITC-conjugated eubacterial cocktail (green) and a Cy3-conjugated *Phaeobacter* (Phaeo.) probe confirmed that many of the bacteria in the jelly capsule were *Phaeobacter* species. The exterior (ext.) and interior (int.) of the capsule are labeled. White arrowheads indicate tubules. Bars, 30µm (a) 50 µm (b&c) and 20µm (d and e).

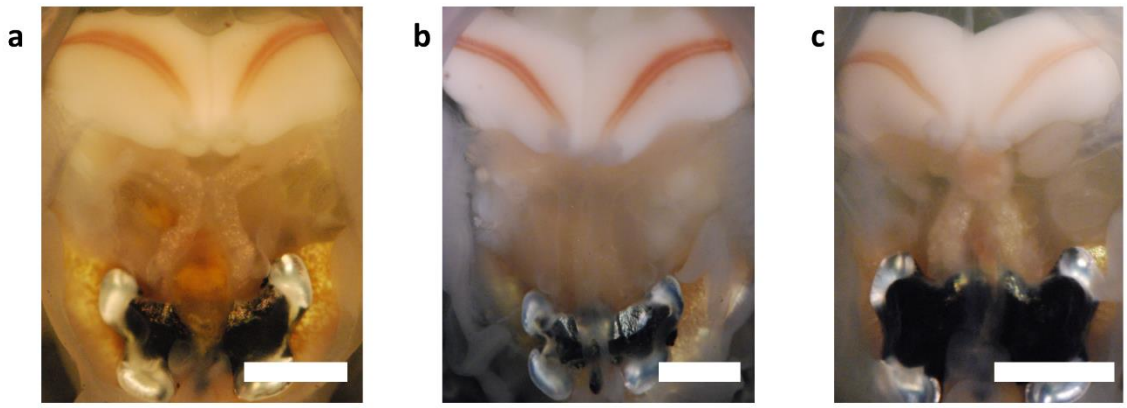

Figure 6. Underdevelopment of ANGs in squid raised in artificial seawater. Animals raised in aquaria containing filer-sterilized seawater and autoclaved sand failed to develop like organs from wild-caught animals. Most animals had no visible ANG (A & B) while some had a severely underdeveloped organ (C). Scale bars = 0.5cm.

Chapter 4. Comparative Genomics of *Roseobacter* Clade Bacteria

Introduction

The *Roseobacter* clade is a vast and diverse group of marine *Alphaproteobacteria.* This group is estimated to account for 10% of all marine bacteria, with higher percentages in coastal seawater (35). These organisms have usually been investigated from an ecological perspective due to their abundance in seawater. The combined metabolic potential of a large population of bacteria has been suggested to contribute to both sulfur cycling, primarily through metabolism of dimethylsulfoniopropionate (DMSP), and carbon cycling as roseobacters oxidize a variety of carbon sources to CO_2 (40).

Many of the characterized isolates of the *Roseobacter* lineage could be described as free-living, but some are associated with other living organisms, including oysters (47), sponges (43) and algae (44; 156). A unique opportunity for researching *Roseobacter*-host associations is the accessory nidamental gland (ANG) of cephalopods (48, 50, 51). The ANG is part of the reproductive tract of female squid and are comprised of many epithelium-lined tubules (50, 121) that house dense populations of bacterial symbionts. These bacteria are secreted into the jelly capsule of the squid's eggs (50, 146). These eggs are deposited in masses on the ocean floor and resist fouling and degradation, even after the embryos inside of them have developed and hatched (118). Studies have investigated the symbionts that inhabit these specialized organs and many have found representatives of the *Roseobacter* clade, including the ANG of *Euprymna scolopes* (48-51). *Roseobacter* clade bacteria, predominantly from the genus *Phaeobacter*, dominate the consortium that also includes *Flavobacteriia* and

Verrucomicrobia. It was also shown that these groups are partitioned into separate tubules, so that only one taxon dominates any given tubule (50).

The bacteria deposited in these jelly coats may play a role in protecting the egg masses from degradation, possibly through the production of antimicrobial compounds. Previous research in the eggs of shrimp have shown that, once the eggs are brooded, *Alteromonas sp.* colonize the surface of the egg and produce the antimicrobial compound 2, 3-indolinedione (157). If these bacteria are removed with antibiotics, the eggs become infected with fungi and do not survive. However shrimp eggs acquire the bacterial epibionts from the seawater. This is an important distinction from squid eggs where the bacterial symbionts from the ANG are actively deposited into jelly-capsule layers by females.

Roseobacter clade organisms are known to produce several antimicrobial compounds, including tropodithietic acid, which has antimicrobial and anti-algal properties (158). Under certain conditions, likely when associated with dying algae, *Phaeobacter gallaeciensis* can produce anti-algal compounds known as roseobacticides derived from p-coumaric acid (159). *Phaeobacter sp.* Y4I and *Phaeobacter daeponensis* produce indigoidine, an antimicrobial blue pigment (148, 160). It is synthesized from a unique polyketide/non-ribosomal peptide synthase gene cluster and has been shown to inhibit marine bacteria, including *Vibrio fischeri* (148).

The function of the ANG and its associated bacterial population remains unknown. Although the bacterial community may produce antimicrobial compounds to provide antifouling protection for the egg masses, no metabolite has ever been identified from *Roseobacter* clade bacteria associated with an ANG or egg-casings (146, 147, 161).

What is unique to these bacterial symbionts? What adaptations have they undergone to survive in a specialized organ such as the ANG? In order to shed light on the metabolic capability of these isolates, we examined the genomes of 12 isolates from the ANG of *E. scolopes* and investigate possible adaptations of these symbionts to their host.

Methods

Culturing bacteria from the ANG

Animals were collected in the sand shallows of Oahu, Hawaii and maintained as previously described (75). To obtain ANGs, mature females were anesthetized in Instant Ocean (Aquarium Systems, Blacksburg, VA) with 2% ethanol. Organs were removed and surface sterilized with 70% ethanol before being homogenized in filter sterilized Squid Ringer's solution (530 mM NaCl, 25 mM MgCl2, 10 mM CaCl2, 20 mM HEPES, pH = 7.5). Homogenate was serially diluted and plated on either salt water tryptone (SWT) or Reasoner's 2A (R2A) medium supplemented with 70% Instant Ocean (34, 125). Plates were incubated aerobically at 28°C for 2-4 days. For each animal, colonies with a unique appearance were isolated for further analysis.

Genome sequencing and annotation

Genomic DNA was isolated from overnight liquid cultures of bacterial isolates using the MasterPure DNA Extraction kit (Epicentre, Madison, WI). DNA was quantified using Qubit fluorescence assay (Life Technologies, Grand Island, NY). Illumina sequencing libraries were created from 1ng of genomic DNA using the Nextera XT kit and the libraries were quantified by HS DNA Bioanalyzer assay (Agilent, Santa Clara, CA). Libraries were sequenced on an Illumina MiSeq sequencer using 2 x 250 bp reads. For *Phaeobacter sp.* ANGI, additional data was added from a previous

sequencing effort using an Illumina mated-pair library (144). Genomes were assembled using the CLC Genomic Workbench (CLC) using default parameters. Assemblies annotated using the Rapid Annotation using Subsytem Technology server for gene calling (162, rast.nmpdr.org). Genomes were also analyzed with Anti-SMASH (Antibiotic and Secondary Metabolite Analysis Shell, 163, antismash.secondarymetabolites.org) and BAGEL3 (BActeriocin Genome mining tool, 164, bagel.molgenrug.nl) for secondary metabolite and bacteriocin gene clusters.

Taxonomic analysis

To ensure equal gene calling across the genomes, all genomes, including the 57 draft and completed genomes obtained from the NCBI ftp, were re-annotated using the RAST server. Assembled contigs were reconstructed from the RAST-generated GenBank files for all genomes using the seqret application of the emboss package (165) .

Phylogenetic methodology

A total of seventy-nine genomes were used in this study. Fifty-seven *Roseobacter* genomes were obtained from the NCBI ftp site (ftp://ftp.ncbi.nih.gov/genomes/) and twelve are new to this study, including an improved assembly of *Phaeobacter sp.* ANG1. (Table 1). Ten non-*Roseobacter* genomes were used as outgroups for taxonomic analysis. Thirty-three housekeeping genes were selected from among the seventy conserved single-copy genes used in a previous phylogeny of the *Roseobacter* lineage (166). DNA queries were obtained from the completed genome of *Roseobacter litoralis* Och 149. The top hits for each gene were aligned separately using MUSCLE (167) and evaluated by hand to verify that the sequences were homologs. In house python scripts created a concatenated alignment from all 33 genes. An optimal model of evolution was determined using the

Akaike Information Criterion with correction for small sample size (AICc). The program jModelTest 2.1.4 was used to compute likelihoods from the nucleotide alignment and to perform the AICc (168, 169). The best-fitting model reported was GTR + Gamma estimation + Invariable site estimation. A maximum likelihood (ML) phylogeny was generated from the concatenated multi-sequence alignment using the PhyML v3.0_360-500M (169). PhyML parameters consisted of GTR model, estimated p-invar, 4 substitution rate categories, estimated gamma distribution, subtree pruning and regrafting enabled with 100 bootstrap replicates. The resulting tree was rooted on the split between Clade 1 and 2, as described by Newton and colleagues (166) and the remainder of the taxa (Clades, 3, 4 and 5).

Average Nucleotide Identity

JSpecies1.2.1 (170) analyzed the genomes for Average Nucleotide Identity (ANI) and tetramer frequency patterns. As the relationships of interest for this study are primarily among presumably closely related species and genera, only the nucmer and tetra algorithms were used. The Blast-based ANI was not used as we were primarily interested in understanding the degree of relatedness between closely related organisms, which the nucmer method is equally capable of (170). Additionally, the increased rate of dropoff between moderately divergent sequences (<90%) the nucmer method yields relative to the Blast method (170) was useful in highlighting when organisms were not highly similar. The default settings for both algorithms were used (170).

Siderophore biochemical assays

In these experiments, levels of contaminating iron were controlled by using water treated with a Nanopure Diamond filtration system (Barnstead, Lake Balboa, CA) to

wash all glassware and make all media components. Siderophore production was confirmed using chrome azurol S (CAS) agar, modified for marine bacteria as previously described (171).

To test survival of ANG isolates in iron limiting conditions, growth in the presence of the iron chelator ethylenediamine-N,N'-bis (2-hydroxyphenylacetic acid) (EDDHA) was examined as described previously (172). Cultures were grown for 24 hours at 26°C in SWT then washed 3x in minimal seasalts solution (MSS, 50 mM $MgSO_4$, 10 mM $CaCl_2$, 350 mM NaCl, 10 mM KCl, 18.5 mM NH_4Cl, 333 µM K_2PO_4, $FeCl_3$ 10 µM, 100 mM PIPES, pH =7.2). Cultures were inoculated to an OD_{600} of 0.05 in MSS with iron added to 10µM. To create iron limiting conditions, EDDHA was added at 0 - 30 µM to limit the availability of free iron. Glucose and casamino acids were added as carbon sources at 0.2% and 0.3%, respectively. After cultures were allowed to grow for 24 h at 26°C, the OD_{600} of cultures with EDDHA added were compared with control cultures without the iron chelator. Siderophore production was measured using the CAS liquid assay as described previously (173). Further chemical characterization of siderophores was done using the Arnow and Csaky assays (174, 175).

Homoserine lactone detection

Homoserine lactone (HSL) production was detected using the HSL sensing strain *Agrobacterium tumefaciens* NTL4 (pZLR4) (176). To determine AHL production, we used a well-diffusion assay as previously described (177). Briefly, a 3 mL culture of *A. tumefaciens* NTL4 was grown for 24 h in LB with gentamicin added to 30µg/mL at 28°C. One milliliter of this culture was used to inoculate 50 mL of AB minimal media (178) containing 0.5% glucose and 0.5% casamino acids. After 24 h incubation, 100 mL of AB

minimal media containing 1.2% agar was autoclaved. Once the molten agar had cooled sufficiently, glucose and casamino acids were added to 0.5% each and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) was added to 75 µg/mL. The molten agar was then combined with the 24 h culture of *A. tumefaciens* and evenly distributed into petri dishes and allowed to solidify.

To induce HSL production by ANG isolates, cultures were grown overnight at 26°C in either SWT or MSS (see above) with $FeCl_3$ added to 30 µM and glucose and casamino acids were added as carbon sources at 0.5% each. In order to prevent the degradation of HSL in alkaline conditions, the growth media was buffered to pH 6.8. In our experiments, the pH of the growth media never rose above 7.5. After a 24 h incubation, the cells were pelleted by centrifugation and the supernatant was filtered through a 0.22 micron filter. Wells were created into the *A. tumefaciens* agar plates and 60µl of cell-free supernatant was deposited into each well.

ANG tissue was tested for the presence of AHLs by dissecting an ANG from a mature female as described above. The ANG was homogenized in 300µL of Squid Ringer's solution and the homogenate was centrifuged at 1,000 x g for 10 min. The supernatant was removed and centrifuged again at 10,000 x g for 10 minutes and 60µl of the resulting clarified homogenate was deposited in a well of the AHL detection plates. To control for any non-specific reaction with metabolites supplied by the host tissue, a gill was also homogenized. The supernatant was clarified as described for ANG tissue and also tested with AHL detection agar. All AHL detection plates were incubated at 28°C and photographed after 48 hours of incubation.

Results and Discussion

Taxonomic analysis

Average nucleotide identity (ANI) suggests that the 12 sequenced isolates can be putatively classified into 7 different species using the accepted ANI cut off of 95% (Figure 1, 170). Assigning a preliminary taxa by top blast hit of the 16S gene, there are two *Ruegeria* species (R and S4), two *Nautella* species (Vp and M1), a single *Tateyamaria* species (S1) and two species of *Phaeobacter*. Previous research has shown that *Phaeobacter* species are the most common symbiont within the ANG (50, 144). One species of *Phaeobacter* was consistently isolated from the four different female organs used in this study. This cluster of isolates likely represents the dominant symbiont present in the ANG.

Most of the isolates from the ANG, including the dominant *Phaeobacter* species (e.g. ANG1, DT, M6 etc.) cluster in a *Leisingera*/*Phaeobacter* group of the *Roseobacter* clade which include organisms such as *Phaeobacter daeponensis, Phaeobacter caeruleus*, *Phaeobacter sp.* Y4I and two members of the *Leisingera* genus (Figure 2).

Taxonomic analysis of 33 single copy genes from the genome show that most of the isolates that predominate the ANG are located in Clade 1 of the *Roseobacter* clade as described by Newton and colleagues (Figure 2, 166). The exception to this is *Tateyamaria sp*. S1, which is in Clade 2. Clade 1 is still a phylogenetically diverse group, but members of this clade are known to have several ribosomal operons and all of the genes necessary for biotin synthesis. Organisms in Clade 1 lack photosynthetic pigments, indicative of their obligate heterotrophic lifestyles (166).

Genome characteristics and general metabolism.

The sequenced *Roseobacter* genomes were fairly large, especially when compared with genomes of photosynthetic organisms, such as Cyanobacteria and *Pelagibacter ubique* (Table 1; 166). The common *Phaeobacter* isolates had genomes that were approximately 4.6Mb in size and the *Nautella* species were even larger, exceeding 5Mb in size (Table 1). However, this is not that much different from the "typical" roseobacter genome, which averages 4.2 Mb (median = 4.4 Mb). These large genomes reflect the diverse metabolism present in the *Roseobacter* clade. As these organisms are non-photosynthetic, they are able to dominate their ecosystems by being "generalists", using many different transporters and enzymes to import and utilize a diverse set of metabolites (166).

Of the ANG symbionts, all have a complete Entner-Doudoroff pathway for metabolizing glucose. Furthermore, all of them lack the key enzyme phosphofructokinase from the Embden-Meyerhof-Parnas pathway. This is typical of many previously sequenced and complete genomes from the *Roseobacter* lineage (35,166, 179). Two organisms (*Tateyamaria sp*. S1 and *Ruegeria sp.* S4) contain a complete pentose-phosphate pathway. The others lack the enzyme 6-phosphogluconate dehydrogenase. As an alternative metabolic pathway, 6-phosphogluconate produced by the first two enzymes of the pentose phosphate pathway could feed into the Entner-Doudoroff pathway (180, 181).

While the *Roseobacter* clade was originally described as a group of obligately aerobic organisms, recently it has been shown that some members contain enzymes needed for anaerobic respiration of nitrate (160). All of the isolates from the ANG

contain nitrate reductase that could potentially be used for anaerobic respiration. Most isolates, with the exception of *Tateyamaria sp.* S1, also contain denitrifying enzymes to respire nitrous oxides as an alternative to aerobic respiration. These data suggest that the ANG isolates may be able to survive and thrive in anaerobic environments by respiring nitrogenous oxyanions.

Type VI secretion system

An interesting feature of the *Phaeobacter*/*Leisingera* clade is that all sequenced genomes contain a Type VI secretion system (T6SS, Figure 2). In *Phaeobacter daeponensis* and *Phaeobacter caeruleus* it has been shown that this T6SS exists on a plasmid (160, 182). The T6SS of the ANG isolates also appears to exist on a plasmid. In *Phaeobacter sp.* ANG1 the T6SS is located on a large contig (>500 kb) containing RepAB plasmid partitioning genes. Similar conclusions were reached in the genome of *Phaeobacter caeruleus, Phaeobacter daeponensis, Leisingera methylohalidivorans* and *Leisingera aquimarina*. Each of these organisms contains a T6SS on plasmids which can vary greatly in size (from 109 kB in *P. caeruleus* to 526 kB in *Phaeobacter sp.* ANG1) however, all have a DnaA 1-like replicase (160, 183, 184). While other members of the *Roseobacter* clade contain a T6SS, the conservation of the T6SS on similar plasmids could be a characteristic of this *Leisingera*/*Phaeobacter* group.

Several functions of the T6SS have been proposed, including antimicrobial functions, as evidenced by direct cell-contact mediated killing (185, 186). It has also been shown to be involved with host-microbial interactions, particularly in the *Rhizobiales*. *Agrobacterium tumefaciens* shows attenuated ability to create tumors when the T6SS is deleted (187). Similarly, the nitrogen-fixing *Rhizobium leguminosarum*

lacking a T6SS will successfully colonize its host, however it will fail to fix nitrogen (188).

Given the wide distribution of the T6SS in genomes of bacteria isolated from the environment, it is likely that many of the T6SS are not antagonistic, but may be involved with general interactions with other microbes, including bacteria and single cell eukaryotes. For example, a T6SS was shown to aid *V. cholerae* in evading predation by amoebae, but the same T6SS had no effect on virulence (189, 190). In *Proteus mirabilis*, swarming motility can create "boundaries" when two strains meet on an agar plate. These boundaries can be altered by manipulating a homolog of the T6SS protein VgrG (191).

It is interesting that many of the isolates have a T6SS, including isolates outside of the *Phaeobacter/Leisingera* clade. Only the *Ruegeria sp*. R does not have a T6SS. This suggests that the T6SS in these bacteria may be important for communication either with the host or with other bacteria. In the ANG of *E. scolopes*, such high densities of bacterial cells will foster frequent close contact with the host epithelium, hemocytes and other bacteria. Given that it functions by direct cell-to-cell contact, the T6SS would be an ideal system for many of these interactions. The T6SS may be important as to how these organisms are selected from the environment and may explain how some species are able to dominate the bacterial populations within a given tubule (50).

Secondary metabolites

The *Roseobacte*r clade has been shown to produce unique secondary metabolites. Some of the most interesting ones include antibacterials such as tropodithietic acid (TDA) produced by organisms such as *Phaeobacter inhibens* and *Ruegeria sp.* TM1040,

and the blue pigment indigoidine produced by organisms such as *Phaeobacter* sp. Y4I and *Phaeobacter daeponensis* (148, 192). None of the biosynthetic pathways for either of these compounds were found in any of the genomes sequenced. Furthermore, no classical antibiotic synthesis pathways (e.g. tetracycline, carbapenems, etc.) were detected.

Analysis with the Antibiotic and Secondary Metabolite Analysis Shell (AntiSMASH, 163) revealed several gene clusters with potential secondary metabolism (Table 2). These included gene clusters for siderophore synthesis, autoinducer synthases (*luxI*), polyketide/non-ribosomal peptide synthases (PKS/NRPS) and volatile compounds such as terpenes. The BActeriocin GEnome mining tooL (BAGEL, 164), was used to screen genomes for possible bacteriocin producing gene clusters, which were found in the *Ruegeria* isolates (R and S4) as well as *Tateyamaria sp*. S1 (Table 2).

All isolates have a conserved non-ribosomal peptide/polyketide synthase gene cluster characterized earlier (193). This gene cluster is highly conserved in the *Roseobacter* lineage, being found in 28 of 57 previously sequenced genomes, and is comprised of 4 genes: a nonribosomal polypeptide synthase, a polyketide synthase, a glycosyltransferase and a phosphopantetheinyl transferase. However, the product of this gene cluster has not yet been characterized. Given that this gene cluster is well-conserved throughout the *Roseobacter* lineage, its importance should be elucidated through future experiments.

Homoserine Lactones

Homoserine lactones produced by LuxI homologs have been widely studied as quorum sensing molecules in bacteria, including the *luxIR* system of *Vibrio fischeri*, the

light organ symbiont of *E. scolopes* (194). AntiSMASH detected 2 separate *luxIR* homologs in the ANG isolates that were most similar to the *ssaIR* and *ssbIR* previously described in *Ruegeria sp*. KLH11 (Figure 3, 43). However, only the *Ruegeria* isolates (S4 and R) have both pairs of *luxIR* homologs. Most of the ANG bacteria only have homologs of *ssbIR*. In *Ruegeria sp.* KLH11, these two systems work together to control biofilm formation and motility (43). The genes *ssaI* and *ssaR*, have been shown to regulate the change between adherent and planktonic lifestyles. Increased levels of homoserine lactones promote flagellar growth and motility, while lower levels foster biofilm development. The actions of these genes can be indirectly repressed by *ssbIR*. The fact that so many ANG isolates have only the *ssbIR* homologs suggest that there may be a unique function for these quorum sensing genes independent of the *ssaIR* quorum sensing system. In addition to *ssbIR*, the ubiqioutous *luxIR* homologs in the roseobacter genomes from the ANG are also similar to the *raiIR* genes described in *Rhizobium etli* (195). Both ssbIR and raiIR are known to produce 3-hydroxyl homoserine lactone compounds, but *raiIR* has been shown to control growth and nitrogen fixation, not motility (195, 196). While these symbionts do not have nitrogen fixation genes, future studies could determine if the *luxIR* homologs controlled growth, perhaps preventing the symbionts from overpopulating the tubules that they inhabit within the ANG

In order to determine that homoserine lactones were present in the ANG and being produced by the bacterial symbionts, we tested for the presence of AHLs using a semi-quantitative biosensor assay. All isolates that could grow to high density in liquid medium produced detectable homoserine lactones (Figure 3). Species like *Tateyamaria sp.* S1 failed to grow to very high density and failed to produce enough HSL to be

detected by the assay (not shown). ANG homogenate was also tested and created a small zone of β-galactosidase activity around the assay well, suggesting that HSLs are produced in the ANG and contribute to the symbiosis (Figure 3). No enzymatic activity was observed in the negative control using gill tissue (not shown), showing that results we observed were not from a nonspecific reaction with compounds from squid tissue.

While homoserine lactones were detected in both pure culture and ANG homogenate, gene regulation by HSL-based quorum sensing may be different than what has been described for their nearest homologs in *Ruegeria sp.* KLH11. Most ANG isolates, including the dominant *Phaeobacter* species, lack the *ssaIR* homologs directly responsible for the increase of motility in *Ruegeria* sp. KLH11. This suggests there must be a yet undescribed role for the ssb*IR* homologs in the *Roseobacter* clade isolates from the ANG.

Future research should investigate the chemical nature of the HSL produced by the autoinducer synthases in individual ANG isolates. The nearest characterized homologs, both RaiI in *Rhizobium etli* and SsbI in *Ruegeria* sp. KLH11 produce C_8-3-hydroxyl-homoserine lactone compounds (43, 195). Future studies should confirm that this kind of HSL is also produced by the members of the ANG consortium. Once genetics are established with these organisms, creating a non-functioning mutant of the autoinducer synthase could reveal phenotypes controlled by quorum sensing. Transcriptomics between a HSL^- strain and wildtype could also reveal what genes may be controlled at high cell densities. Further studies should also examine how gene expression changes from the high-cell density environment of the ANG to the egg capsule, where cell densities will be lower, but where any anti-fouling compounds may

be produced. Given the profound differences in cell density between free-living symbionts in seawater and the symbionts in the egg capsule and ANG, quorum sensing is an ideal method for gene regulation.

Siderophores

Another group of major secondary metabolites that was detected in the genomes of ANG isolates were siderophores. Siderophores are small molecules with very high affinity for iron. Iron is essential for many cellular functions, including respiration, detoxification of reactive oxygen species (e.g. catalases, superoxide dismutase), and metabolism (e.g. aconitase of the TCA cycle) and very few organisms are known to survive without iron (197). One way that bacteria can acquire iron in iron-limiting conditions is by producing siderophores to bind and sequester iron from other sources.

Siderophore production is unusual in the *Roseobacter* clade. Of the previously sequenced 57 roseobacter genomes, only 6 are predicted to have siderophore synthesis genes (Figure 2). However, all roseobacters isolated from the ANG of *Euprymna scolopes,* with the exception of *Ruegeria sp.* R and S4, have either siderophore biosynthesis genes or showed siderophore activity in biochemical assays. For example, *Nautella sp.* M1 had no predicted siderophore synthesis genes, but siderophore activity was detected when grown on CAS agar, suggesting that these genes may not be annotated, perhaps due to the fragmented state of the assembled genome for this isolate. Conversely, *Tateyamaria sp.* S1 has siderophore biosynthetic genes, but failed to show siderophore activity. Taken together, this suggests induction of siderophore synthesis genes may be controlled very differently in *Tateyamaria sp*. S1 and may be induced only under certain conditions.

To determine how siderophore synthesis might contribute to the lifestyle of the ANG symbionts, we examined growth and siderophore production of *Phaeobacter sp.* ANG1, a representative of the dominant ANG symbionts, against other species from the *Roseobacter* lineage. Siderophore producing strains *Phaeobacter inhibens* DSMZ 17395 and *Leisingera methylohalidivorans* DSM 14336 were tested along with the non-siderophore producing strain *Phaeobacter sp.* Y4I. When grown in the presence of the iron chelator EDDHA, other *Roseobacter* strains had a growth defect, growing to only 20% of the control density. However, *Phaeobacter sp. ANG1* had a much smaller growth defect ($p < 0.001$), growing to greater than 50% of the control OD when concentrations of EDDHA were three times the concentration of available iron in the media (Figure 4a). EDDHA was not toxic to any of our strains, as supplying excess iron to overwhelm the chelator restored growth in all organisms to the same level as the controls with no iron chelator (Figure 4b). The survival of *Phaeobacter sp.* ANG1 under iron-limiting conditions could be due to the higher levels of siderophore produced by this bacterium. Supernatant from cultures of strains that failed to grow (*P. inhibens* and *L. methylohalidivorans*) showed very little CAS activity while supernatants from cultures of ANG1 had significantly higher levels of CAS activity ($p < 0.001$), indicative of a high concentration of siderophores (Figure 4c).

To determine if this higher level of siderophores produced by ANG1 was independent of the growth and survival of this organism, CAS activity was measured in supernatants from cultures without any iron chelator added. This allowed the bacteria to grow and deplete the iron available in the media, leading to induction of siderophore synthesis. Supernatants from cultures of *Phaeobacter sp.* ANG1 had more CAS activity

than either *P. inhibens* DSM17395 or *L. methylohalidivorans* DSM14336 per unit OD_{600} (Figure 4d, $p < 0.01$). This shows that the high amount of siderophores produced by *Phaeobacter sp.* ANG1 is not just due to an increase in cell number, but due to increased siderophore production on the cellular level.

Examining the siderophore biosynthesis genes in roseobacters isolated from the ANG, reveals a unique genome rearrangement. In all other siderophore-producing *Roseobacter* clade organisms, siderophore synthesis genes are located downstream of an iron membrane receptor and an iron-compound ABC transporter. In roseobacters isolated from the ANG, four genes related to polyamine metabolism are inserted upstream of the iron membrane receptor (Figure 5). Polyamines can be used in the production of certain hydroxamate and catechol siderophores (198, 199). These particular polyamine metabolism genes are sufficient to synthesize putrescine, a backbone of certain catechol siderophores such as photobactin from *Photorhabdus luminescens* (200). Testing the supernatant of *Phaeobacter sp.* ANG1 showed siderophores produced by these organisms was also of the catechol type. Future research may determine if the catechol siderophore produced by the ANG symbionts contains putrescine or another polyamine as a structural component. This may help explain the unique genome rearrangement within these organisms. Not only would this insertion change regulatory elements around the siderophore biosynthesis genes, but by coupling polyamine production with siderophore synthesis, there would be sufficient raw materials to efficiently produce more siderophores than other organisms.

Higher levels of siderophore production could be beneficial in colonizing host tissue, which can be an iron-depleted environment due to the host actively sequestering

iron. An infected host can produce more iron-chelating proteins as a way to starve infectious bacteria of a critical resource. One of the most-widely studied examples of this is the reaction of intestinal cells to pathogenic bacteria. The siderophore enterobactin is produced by several species of enteric bacteria, including *Salmonella* and *E. coli* species (201). This siderophore acquires iron from serum proteins carrying iron, such as transferrin, and the siderophore-iron complex is transported into the infecting bacteria to keep them supplied with iron. To combat this, the innate immune system produces proteins to bind siderophores, such as siderocalin/lipocalin, to bind enterobactin and prevent the iron-scavenging molecules from fulfilling their function (202, 203).

Similar mechanisms could be at work in the ANG. The iron chelating proteins ferritin and transferrin have been found in transcriptomes from both hemocytes and light organs of *E. scolopes* (A. J. Collins, unpublished data). These ubiquitous iron proteins primarily can act as iron storage and iron transfer proteins in eukaryotes. Ferritin is present in the hemolymph of invertebrates where it can function as an iron transporter or iron scavenger (204). Transferrin is produced in the epithelium of insects and its expression is up-regulated in response to a bacterial infection (205, 206). These iron chelators could be present in the tubules of the ANG and provide a selective pressure that other roseobacters to overcome. In such a case, siderophore-producing organisms may have an advantage over others, thus explaining why siderophore-producing roseobacters such as *Phaeobacter sp.* ANG1 dominate the consortium. Colonization of the ANG is likely a complex process regulated by many factors. Overcoming iron-limitation by the host is likely just one part of the establishment of the ANG consortium.

This study sets the foundation for future research on the ANG symbionts by characterizing the genomes of several isolates from the *Roseobacter* lineage. We have identified many interesting features in these genomes including Type VI secretion systems, putative quorum sensing systems using homoserine lactones and siderophore production. However, the function of the ANG still remains unknown and it is difficult to understand what role the bacteria play without knowing how the ANG and its bacterial consortium benefits the many cephalopods that have such an organ. There are many exciting avenues for research in this system that will reveal what the function of the consortium is and what adaptations the symbionts have made to thrive in such a specialized organ.

= ANI ≥ 98% + Tetra > 0.9900

= ANI ≥ 95% + Tetra > 0.9900

= ANI ≥ 90% + Tetra > 0.9900

+ Any color = Tetra < 0.9900

Target Genome: / Query Genome:

Query Genome:	*Leisingera aquimarina* DSM 24565	*Leisingera methylohalidivorans* DSM 14336	*Nautella sp.* M1	*Nautella sp.* Vp	*Phaeobacter arcticus* DSM 23566	*Phaeobacter caeruleus* DSM 24564	*Phaeobacter daeponensis* DSM 23529	*Phaeobacter gallaeciensis* DSM 26640	*Phaeobacter inhibens* 2.10	*Phaeobacter inhibens* DSM 17395	*Phaeobacter inhibens* DSM 16374	*Roseobacter sp.* MED193	*Roseobacter sp.* SK209-2-6	*Phaeobacter sp.* ANG1	*Phaeobacter sp.* M6	*Phaeobacter sp.* DT	*Phaeobacter sp.* S	*Phaeobacter sp.* S3	*Phaeobacter sp.* S5	*Phaeobacter sp.* M7	*Ruegeria sp.* KLH11	*Ruegeria sp.* R	*Ruegeria sp.* R11	*Ruegeria sp.* S4	*Ruegeria sp.* TW15	*Ruegeria pomeroyi* DSS-3	*Ruegeria sp.* TM1040	*Ruegeria sp.* Trich CH4B	*Ruegeria* ITI-1157	*Tateyamaria sp.* S1
Leisingera aquimarina DSM 24565	100	88.40	85.85	85.96	83.34	86.48	86.26	84.42	84.45	84.54	84.28	83.33	84.09	86.19	86.19	86.22	86.19	86.18	86.29	86.01	83.28	83.55	84.51	83.69	83.61	83.43	83.77	83.65	83.47	82.89
Leisingera methylohalidivorans DSM 14336	88.41	100	85.76	85.76	83.42	86.50	86.25	84.27	84.37	84.30	84.22	83.43	84.22	86.00	86.03	86.04	86.00	85.99	86.02	85.86	83.44	83.26	84.51	83.62	83.27	83.46	83.74	83.71	83.56	82.99
Nautella sp. M1	85.89	85.76	100	89.63	83.34	86.36	86.25	84.55	84.31	84.30	84.42	83.39	84.84	86.30	86.35	86.32	86.31	86.29	86.28	86.15	83.71	83.49	84.91	83.97	83.82	83.69	84.17	84.34	84.06	83.42
Nautella sp. Vp	85.98	85.75	89.62	100	83.35	86.39	86.28	84.63	84.48	84.36	84.52	83.36	84.84	86.35	86.41	86.42	86.36	86.34	86.32	86.27	83.74	83.64	84.87	83.95	83.70	83.63	84.20	84.26	84.06	83.55
Phaeobacter arcticus DSM 23566	83.37	83.42	83.24	83.23	100	83.16	83.76	85.00	83.38	83.35	83.60	85.94	83.71	83.19	83.13	83.11	83.19	83.18	83.17	83.26	82.82	82.78	83.39	83.00	83.08	82.65	83.41	84.11	82.73	82.44
Phaeobacter caeruleus DSM 24564	86.54	86.49	86.37	86.40	83.28	100	90.57	84.45	84.48	84.43	84.45	83.41	84.67	87.43	87.43	87.46	87.42	87.41	87.83	87.14	83.74	83.82	84.89	84.00	83.59	83.84	84.21	84.33	84.12	83.33
Phaeobacter daeponensis DSM 23529	86.29	86.25	86.27	86.30	83.85	90.57	100	84.51	84.35	84.41	84.43	83.45	84.46	87.47	87.47	87.50	87.46	87.54	87.50	87.15	83.62	83.88	84.99	83.83	83.69	84.33	84.29	84.72	84.25	83.27
Phaeobacter gallaeciensis DSM 26640	84.55	84.27	84.56	84.64	85.15	84.42	84.50	100	89.84	89.98	89.92	85.14	84.57	84.67	84.67	84.72	84.68	84.68	84.62	84.40	83.83	83.84	85.09	83.75	83.89	83.21	84.31	84.50	83.91	83.56
Phaeobacter inhibens 2.10	84.55	84.38	84.33	84.49	83.58	84.48	84.35	89.84	100	97.18	97.69	83.66	84.58	84.63	84.74	84.6	84.65	84.65	84.68	84.37	83.72	83.78	84.97	83.90	83.63	83.39	84.44	84.39	83.88	83.42
Phaeobacter inhibens DSM 17395	84.66	84.29	84.30	84.37	83.54	84.44	84.39	89.98	97.19	100	97.71	83.60	84.54	84.66	84.76	84.66	84.69	84.76	84.57	84.35	83.61	83.69	85.01	83.89	83.88	83.19	84.17	84.44	83.99	83.38
Phaeobacter inhibens DSM 16374	84.40	84.24	84.43	84.54	83.82	84.45	84.42	89.92	97.67	97.69	100	83.87	84.54	84.64	84.66	84.53	84.66	84.65	84.59	84.49	83.62	83.75	85.03	83.90	83.82	83.25	84.35	84.36	83.88	83.58
Roseobacter sp. MED193	83.57	83.43	83.40	83.38	85.98	83.40	83.45	85.13	83.66	83.60	83.84	100	84.33	83.42	83.45	83.40	83.46	83.45	83.44	83.37	83.39	83.32	83.64	83.76	83.62	83.01	83.85	84.34	83.47	83.28
Roseobacter sp. SK209-2-6	84.12	84.22	84.57	84.57	83.73	84.24	84.12	84.11	83.99	83.92	83.91	83.91	100	84.55	84.51	84.52	84.56	84.55	84.48	84.44	83.56	83.66	84.37	83.77	83.45	82.95	83.51	83.52	83.25	82.85
Phaeobacter sp. ANG1	86.21	86.00	86.3	86.35	83.32	87.42	87.46	84.69	84.65	84.68	84.65	83.41	84.91	100	97.92	97.88	99.93	99.90	97.79	90.11	83.81	83.59	85.17	83.97	83.82	83.87	84.19	84.12	84.21	83.49
Phaeobacter sp. M6	86.21	86.04	86.36	86.41	83.26	87.42	87.47	84.67	84.75	84.77	84.66	83.44	84.82	97.94	100	97.88	97.93	97.95	97.89	90.10	83.73	83.67	85.15	83.99	83.97	83.79	84.08	84.22	84.23	83.56
Phaeobacter sp. DT	86.25	86.03	86.32	86.42	83.24	87.45	87.48	84.70	84.58	84.66	84.53	83.39	84.87	97.88	97.86	100	97.87	97.86	97.85	90.11	83.75	83.79	85.23	84.06	83.92	83.85	84.13	84.16	84.21	83.49
Phaeobacter sp. S	86.21	86.00	86.31	86.36	83.33	87.42	87.46	84.68	84.63	84.69	84.64	83.46	84.91	99.96	97.92	97.88	100	99.91	97.80	90.10	83.79	83.59	85.16	83.98	83.82	83.88	84.14	84.08	84.21	83.46
Phaeobacter sp. S3	86.21	85.98	86.3	86.35	83.31	87.41	87.54	84.67	84.63	84.76	84.64	83.46	84.90	99.93	97.95	97.87	99.92	100	97.80	90.09	83.81	83.62	85.19	84.01	83.81	83.89	84.17	84.09	84.24	83.47
Phaeobacter sp. S5	86.31	86.01	86.29	86.33	83.28	87.84	87.50	84.61	84.66	84.57	84.58	83.44	84.81	97.81	97.88	97.86	97.80	97.80	100	90.09	83.80	83.78	85.28	84.04	83.91	83.79	84.15	84.23	84.20	83.65
Phaeobacter sp. M7	86.03	85.85	86.15	86.26	83.41	87.12	87.14	84.39	84.36	84.35	84.48	83.36	84.81	90.11	90.10	90.11	90.10	90.09	90.09	100	83.72	83.74	85.16	84.09	83.87	83.69	84.08	84.26	84.15	83.43
Ruegeria sp. KLH11	83.55	83.40	83.74	83.78	83.28	83.71	83.66	83.87	83.78	83.59	83.66	83.23	84.59	83.81	83.76	83.77	83.81	83.83	83.81	83.74	100	84.50	83.74	84.41	84.24	83.49	84.20	84.76	84.02	83.35
Ruegeria sp. R	83.88	83.25	83.53	83.64	83.26	83.81	83.87	83.83	83.77	83.69	83.72	83.33	84.96	83.55	83.65	83.78	83.60	83.62	83.78	83.74	84.50	100	83.77	84.17	85.11	83.54	84.02	85.04	84.02	83.35
Ruegeria sp. R11	84.65	84.50	84.93	84.88	83.56	84.87	84.99	85.09	84.97	85.01	85.04	83.64	84.92	85.14	85.16	85.22	85.16	85.19	85.28	85.15	83.69	83.78	100	84.2	84.04	83.38	84.56	84.56	83.83	83.54
Ruegeria sp. S4	84.10	83.62	83.97	83.94	83.49	83.94	83.84	83.74	83.91	83.9	83.91	83.76	85.28	83.93	83.98	84.07	83.97	83.99	84.03	84.07	84.41	84.18	84.21	100	84.32	83.73	84.38	84.95	84.03	83.47
Ruegeria sp. TW15	84.04	83.26	83.83	83.72	83.67	83.58	83.69	83.86	83.61	83.87	83.81	83.60	84.89	83.75	83.95	83.93	83.84	83.82	83.94	83.87	84.24	85.11	84.02	84.30	100	83.28	84.53	84.80	83.90	83.58
Ruegeria pomeroyi DSS-3	83.58	83.47	83.70	83.63	82.85	83.84	84.33	83.22	83.39	83.19	83.24	83.00	83.68	83.87	83.79	83.86	83.89	83.89	83.82	83.69	83.51	83.56	83.39	83.77	83.29	100	83.51	83.84	84.31	83.04
Ruegeria sp. TM1040	83.79	83.73	83.85	83.93	83.39	83.90	83.99	83.95	84.02	83.77	83.95	83.24	83.55	83.91	83.85	83.92	83.93	83.92	83.90	83.74	83.43	83.26	84.21	83.54	83.63	83.12	100	85.96	83.26	82.90
Ruegeria sp. Trich CH4B	83.70	83.69	83.85	83.81	84.10	83.89	84.42	83.94	83.82	83.84	83.79	83.51	83.58	83.58	83.74	83.69	83.61	83.60	83.75	83.8	83.71	83.75	84.02	83.66	83.53	83.17	85.85	100	83.51	82.84
Ruegeria ITI-1157	83.46	83.54	83.80	83.81	82.73	83.88	84.04	83.38	83.43	83.54	83.43	82.73	83.23	83.96	83.97	83.96	83.97	83.99	83.95	83.89	83.90	83.86	83.46	83.81	83.69	84.15	83.26	83.39	100	83.02
Tateyamaria sp. S1	83.34	82.99	83.43	83.55	83.24	83.29	83.27	83.56	83.42	83.39	83.58	83.27	84.53	83.40	83.54	83.47	83.47	83.48	83.61	83.43	83.31	83.35	83.53	83.51	83.59	83.01	83.91	84.29	83.57	100

Figure 1. Average nucleotide identity (ANI) comparison shows 7 unique isolates from the ANG. Six of the isolates cultured from the ANG are very similar, sharing >99% ANI (DT, ANG1, M6, S, S3 and S5). These species dominate the culturable isolates of the ANG and have been consistently isolated from many different animals. Isolate M7 may be part of this genus, but is distinct from this group. The two *Nautella* species (VP and M1) share nearly 90% ANI and may be part of the same genus. The other isolates, R, S4 and S1 are unique from other sequenced roseobacters. Isolates are color-coded to indicate genus. Green – *Phaeobacter*, Blue – *Ruegeria*, Light blue – *Tateyamaria*, Teal – *Nautella*. (Analysis by Matt Fullmer)

Table 1. Genome assembly statistics for *Roseobacter* clade ANG isolates

Isolate	Genome Size (Mb)	# of genes	%GC	N50 (kb)	Contigs	Coverage
Nautella sp. VP	5.150	4,941	62.3	70	165	69.2
Nautella sp. M1	5.375	5,097	62.0	211	180	132.3
Phaeobacter sp. ANG1	4.587	4,484	62.8	450	36	1,455
Phaeobacter sp. DT	4.596	4,467	62.6	189	116	115.4
Phaeobacter sp. S	4.572	4,458	62.8	196	83	65.5
Phaeobacter sp. S3	4.597	4,468	62.7	300	84	129.0
Phaeobacter sp. M6	4.542	4,429	62.7	157	65	118.0
Phaeobacter sp. S5	4.660	4,534	62.5	233	54	123.5
Phaeobacter sp. M7	4.582	4,498	62.5	263	61	148.7
Ruegeria sp. R	4.685	4,755	57.4	390	47	98.1
Ruegeria sp. S4	4.538	4,619	57.2	978	20	71.9
Tateyamaria sp. S1	4.425	4,478	60.6	229	33	110.7

**Phaeobacter sp.* ANG1 was previously sequenced with an Illumina mate-pair library and is therefore and a much higher fold coverage than other genomes

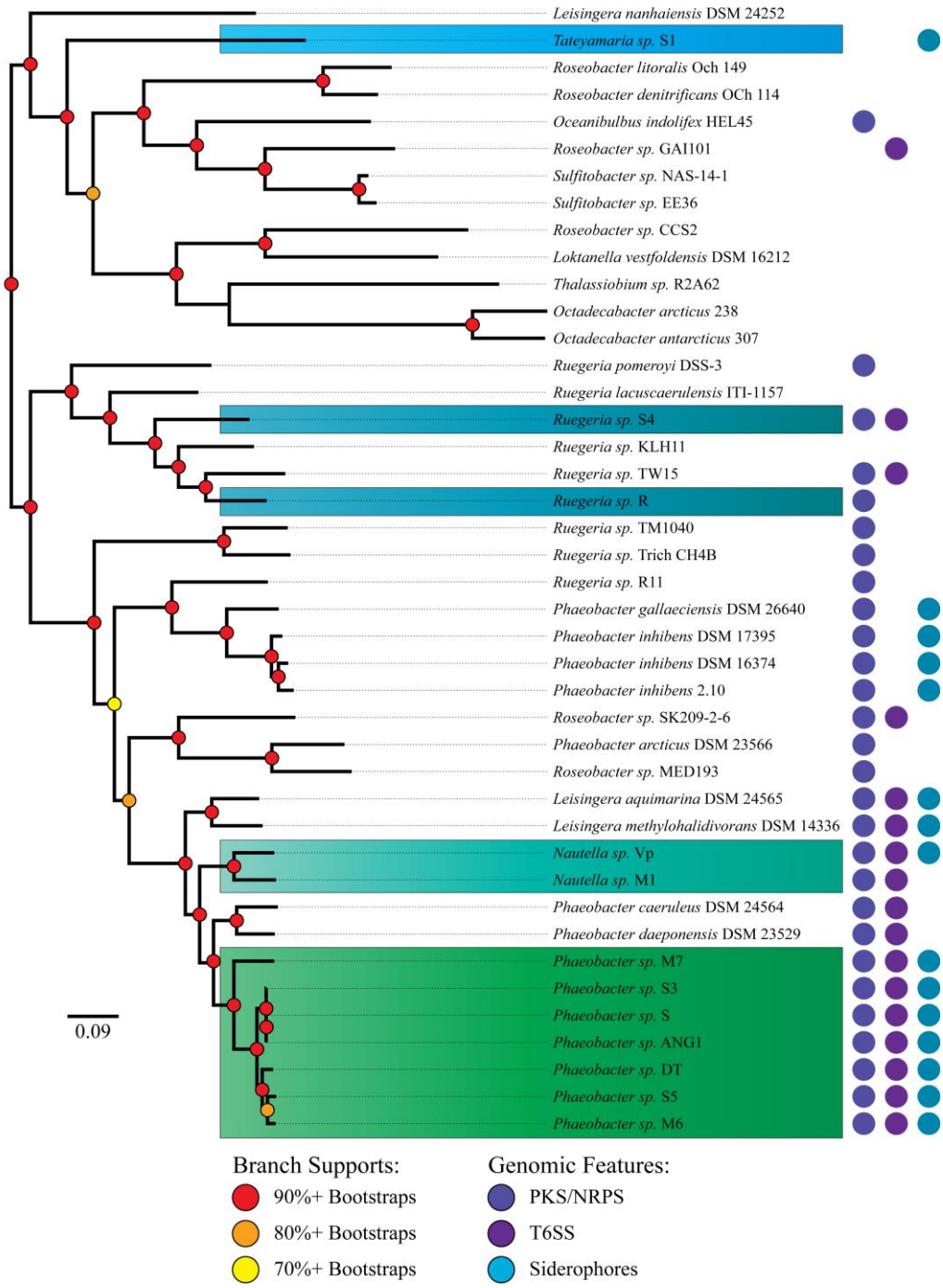

Figure 2. MLSA analysis of *Roseobacter* clade isolates from the ANG with closely-related organisms and distribution of significant gene clusters. Phylogenetic analysis of 33 single-copy housekeeping genes proteins places most ANG isolates in the previously described in "Clade 1" of the *Roseobacter* clade (166). A polyketide/non-ribosomal peptide synthase gene cluster is spread throughout most of the *Roseobacter* clade, while Type VI secretion systems and siderophores are limited to only a few. (Analysis by Matt Fullmer)

Table 2. Secondary metabolites detected with AntiSMASH and BAGEL

	M1	VP	ANG1	S	S3	S5	M6	DT	M7	R	S4	S1
PKS/NRPS	1	1	1	1	1	1	1	1	1	1	2	0
LuxRI	1	1	1	1	1	1	1	1	1	2	2	1
Bacteriocin	0	0	0	0	0	0	0	0	0	3	3	2
Siderophore	0	1	1	1	1	1	1	1	1	0	0	1
Terpene	0	0	0	0	0	0	0	0	0	0	0	1
Ectoine	1	1	0	0	0	0	0	0	0	1	0	0

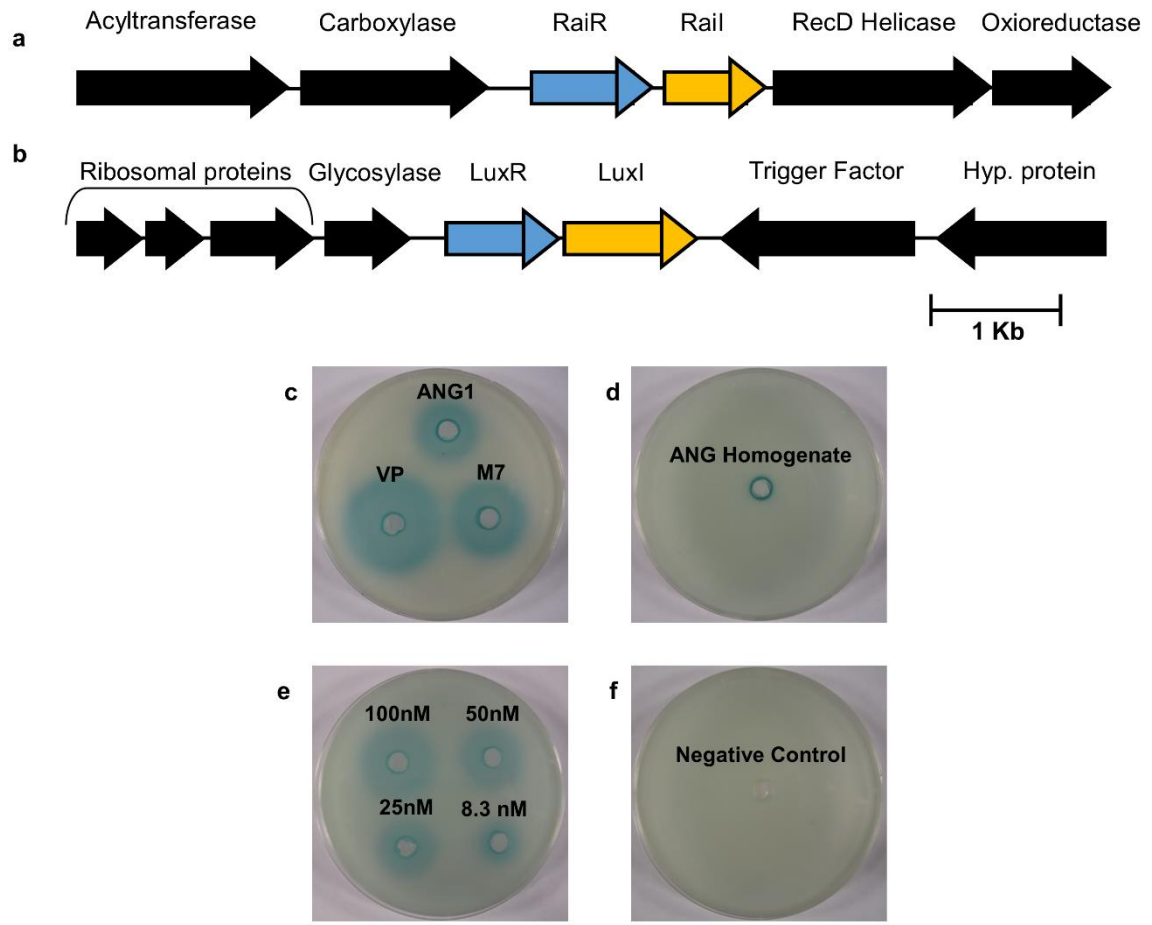

Figure 3. LuxIR homologs in *Roseobacter* clade organisms from the ANG and associated homoserine lactone production. (a) All ANG isolates have a pair of *luxIR* homologs similar to *raiI and raiR* flanked by potential anabolic genes (Crontonyl CoA carboxylase and Acetyltransferase) and a helicase and oxioreductase. (b) Another pair of *luxIR* homologs is present only in the *Ruegeria* species (isolates S4 and R) flanked by ribosomal proteins, a cell division trigger factor and a glycosylase. (c) Representative picture of B-galactosidase activity induced by homoserine lactone in supernatants from cultures of *Phaeobacter sp.* ANG1, *Phaeobacter sp.* M7 and *Nautella sp.* VP. (d) Homoserine lactones were also detected in homogenate of ANG tissue. (e) Semi-quantitative dilution of N-3-oxo-hexanoyl-homoserine lactone. 60µL of each concentration were added to the indicated well. (f) 60µL growth medium (Negative control).

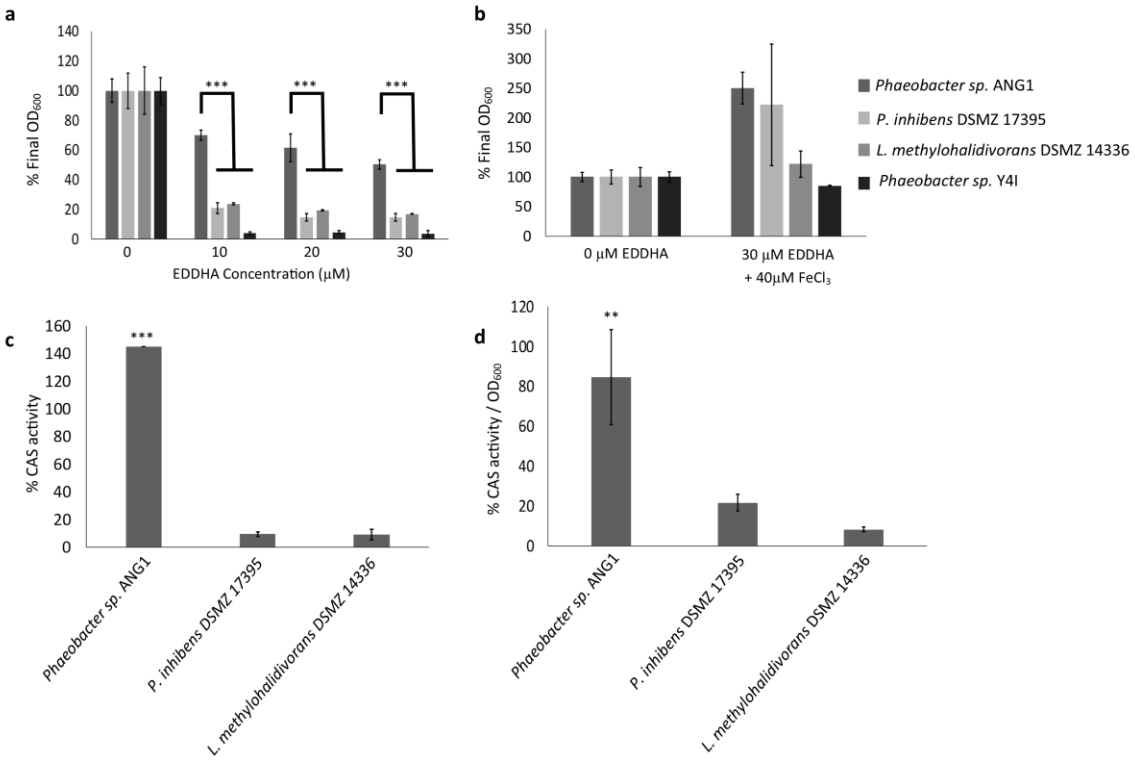

Figure 4. *Roseobacter* clade organisms from the ANG have a growth advantage in iron limiting conditions, possibly due to siderophore production. (a) While other *Roseobacter* clade organisms are inhibited by the presence of an iron-chelator, *Phaeobacter sp*. ANG1 grows to more than 50% of its optical density even if EDDHA is at three times the concentration of available iron. (b) EDDHA is not toxic to *Roseobacter* clade organisms, as adding enough iron to overwhelm the chelator restores the growth defect to all *Roseobacter* clade organisms. (c) ANG isolates produce more siderophore than non-ANG isolates in the presence of EDDHA. (d) Even if no iron chelator is present, siderophores are still present in relatively high amounts in the supernatants from cultures of *Phaeobacter sp.* ANG1. ($p < 0.0001$ indicated by ***, $p < 0.01$ indicated by **)

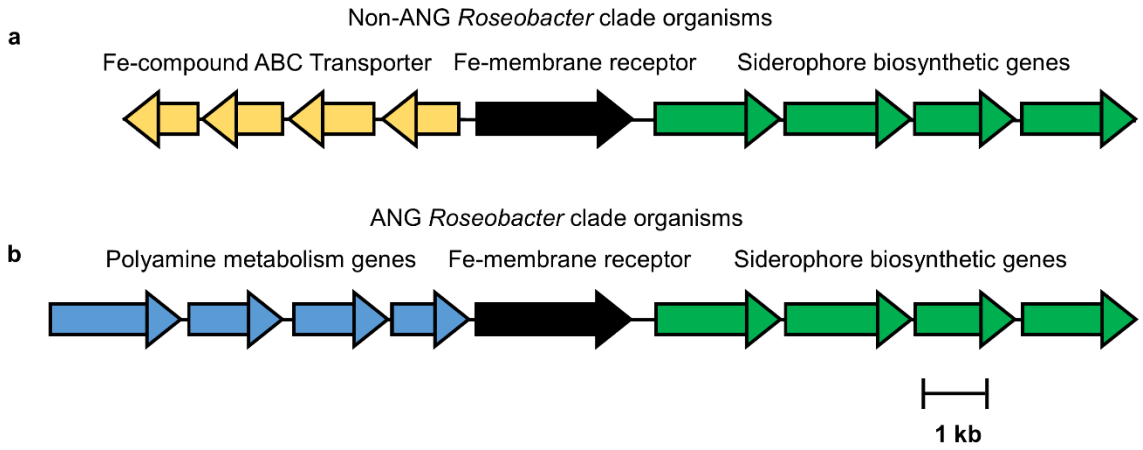

Figure 5. *Roseobacter* isolates from the ANG have a unique genome rearrangement upstream of siderophore biosynthesis group. (a) In previously sequenced *Leisingera* and *Phaeobacter* species, an ABC transporter, predicted to transport iron compounds like iron dictate, lies upstream of a membrane receptor protein and the siderophore synthesis gene cluster. (b) In ANG bacteria, polyamine metabolism genes replace the ABC transporter genes. In addition to modifying regulatory elements, the four polyamine genes could synthesize a polyamine backbone of the siderophores, such as putrescine.

Chapter Five. Future Directions and Conclusions

The light organ symbiosis has been a model for the study of interactions between animals and their bacterial symbionts for more than 20 years (11). Much research has been dedicated to the initiation of the symbiosis and how the two partners, *E. scolopes* and *V. fischeri* communicate and influence one another. But the mechanisms behind how the immune system controls the bacterial symbiont population were unknown. The highly conserved Nf-κB pathway had been detected in light organ tissues from juveniles and formed the preliminary model of molecular mechanisms in hemocytes (89). However, it was not known if any of these genes were expressed in hemocytes. To answer this question, I used high-throughput sequencing to determine which genes were expressed in hemocytes from *E. scolopes*. This provided the first snapshot of genes expressed in hemocytes circulating within *E. scolopes*.

Several members of the NF-κB signaling pathway were found within the hemocyte transcriptome. Including the central regulating protein, NF-κB, along with two other proteins, an activator and an inhibitor of the NF-κB transcriptional protein. NF-κB, however, was not detected in the proteome of circulating hemocytes.

One key protein missing from the transcriptome was the previously described toll-like receptor (TLR). The Toll-like receptors are a diverse group that have been implied in many immune cell functions, including detecting pathogens and the production of many immune effectors, including cytokines (207-209). No other TLRs were discovered which is surprising given the variety of toll-like receptors and their association with macrophage cells. Proteomic methods did detect the previously described Toll-like receptor, but it was in very low abundance (T.R. Schleicher, personal communication).

Further sequencing will hopefully discover more TLRs as well with other PRRs that could be expressed in hemocytes.

Many novel proteins were discovered which need to be further characterized. One very interesting protein is galectin, a sugar binding lectin. Galectin is a pattern recognition receptor that has been shown to differentially bind sugar moieties on certain vibrios (210). Tyler Schleicher has been characterizing the expression and localization of this protein and how it may be important to how hemocytes phagocytose bacteria.

The most exciting protein discovered in the hemocyte transcriptome was a novel PRR, EsPGRP5. This peptidoglycan recognition protein appears to be the most highly expressed PRR in hemocytes as indicated by both qRT-PCR and sequencing coverage from the transcriptome. A short-form PGRP, it is likely to have T7-like amidase activity which cleaves the muramyl peptide from monomers of peptidoglycan. Comparing this protein with other PGRPs from *E. scolopes*, it is the most divergent protein yet described. This could be reflective of its role in a systemic tissue (hemocytes) as compared to the PGRPs described from the light organ, which originate from tissues in direct contact with one bacterium, *V. fischeri*. Much more research needs to be done on this protein, including biochemical characterization of the protein. I pioneered two approaches for expressing this protein *in vitro* (Appendix B) and while recombinant EsPGRP5 was expressed in both methods, I was not able to purify the protein from raw cell lysate. Hopefully future researchers will optimize my techniques and thoroughly characterize the biochemical aspects of EsPGPR5.

In hemocytes, EsPGRP5 protein is located to vacuoles which seem to be closely associated with the actin cytoskeleton. This localization does not change when

hemocytes are challenged with vibrios, such as *V. fischeri* or *V. harveyi*. Some PGRPs are known to be stored in vacuoles and associate with the phagosome after bacteria have been engulfed by immune cells. These PGRPs can affect the survival of phagocytosed Gram positive cells (211). It should be determined if EsPGRP5 co-localizes with phagosomes within the hemocytes and serves a similar function. However, many short form PGRPs are also secreted. While the EsPGRP5 detected by Western blotting has no signal peptide for secretion, future experiments should test for extracellular EsPGRP5, testing both the hemolymph and the crypt spaces of the light organ. Even without a signal peptide, EsPGRP5 could still be secreted by exocytosis. If EsPGRP5 is found in the crypt spaces, it could work alongside EsPGRP2 to detoxify TCT in the light organ.

This work sets the foundation to discover the mechanisms by which hemocytes are able to phagocytose bacteria at different rates. Despite the discovery of many signaling and receptor proteins, how these components work together is still unknown. Future research must look at how these components interact with one another and dictate behavior of the squid. Interfering RNA (RNAi) will no doubt be an invaluable tool. Genetic knock-down of key signaling proteins or PRRs will likely show the precise mechanism by which hemocytes bind and phagocytose bacteria. Biochemical characterization of surface proteins that bind or interact with bacteria and determine whether a cell is phagocytosed should be given priority.

Further work should also examine the maturation of hemocytes in *E. scolopes*. A change in *V. fischeri* binding is seen over several days of antibiotic treatment, which suggests a turnover in hemocyte population. Hemocytes develop in the white body before entering the circulatory system. By what mechanism does removing the

symbionts from the light organ cause a change in hemocyte development, presumably within the white body? The possibility of a distant effect of the light organ symbionts on the white body is a tantalizing question for future research.

At the beginning of my doctoral work, the light organ symbiosis had been a well-characterized system, but little was known about the accessory nidamental gland in *E. scolopes*. It was known from previous studies that *Roseobacter sp.* (as the *Phaeobacter* genus had not been widely accepted at the time) inhabited the accessory nidamental glands of other cephalopods, including *Doryteuthis* and *Sepia* species. These roseobacters were known to be deposited into the egg capsules of freshly laid squid. In fact, it was believed that these organisms could become pathogenic after a few weeks of incubation (212). Egg capsules of *E. scolopes* sporadically turn pink and this was correlated with a decrease in juvenile viability. This was attributed to an infection or bloom of *Roseobacter* species within the egg capsule. It was under this assumption that we began to look at the bacterial population of the ANG.

The ANG in *E. scolopes* is a very specialized organ and a unique niche for *Roseobacter* clade bacteria. Microscopy revealed many epithelium-lined tubules house dense populations of bacteria. Blood vessels were interspersed between the tubules showing that the ANG was highly vascularized. Two major types of epithelium were observed, each containing a unique morphology of bacteria, either a larger coccoid bacterium or a smaller bacillus bacterium.

In order to identify the bacterial symbionts in the ANG, we created and analyzed 16S clone libraries. This revealed that across many individuals, the bacterial symbionts of the ANG are conserved, and dominated by members of the *Roseobacter* lineage, which

includes genera such as *Phaeobacter*, *Ruegeria*, and *Nautella*. Another conserved group of bacteria were the *Verrucomicrobia*, a phylum that remains uncharacterized due to the difficulty of growing isolates in culture. Species from the *Rhizobiales* were also detected in the clone libraries. These bacterial taxa were also confirmed using 454 metagenomic data. This showed the presence of another taxa, present in the ANG *Flavobacteriia*.

Using this taxonomic data, I used 16S rRNA FISH to further examine the population structure of the ANG. Observing the resulting fluorescent signals showed that each tubule was dominated by one bacterial taxon. For example, tubules were seen with only fluorescent signal from *Verrucomicrobia* or *Alphaproteobacteria*. Some tubules were observed to be dominated with 16S probes specific to *Phaeobacter*, suggesting that the bacteria may be partitioned at the genus level.

Future research on the ANG should examine the mechanism of this partitioning. The different types of epithelium in ANG tubules could create a different microenvironment that selects for specific taxon. Many different factors could lead to the dominance of one bacterium over others, including oxygen availability, pH and nutrients provided by the host. Some characteristics, such as pH and oxygen, may be determined by direct measurement of the microenvironment using microsensors. Determining carbon sources used by bacteria in the ANG may be determined indirectly by using transcriptomics. Key pathways for nutrient acquisition should be up-regulated in the ANG and gene expression data can be used to see which carbon sources are being used by the symbionts.

Other host effectors may contribute to the partitioning of bacterial species. The enzyme halide peroxidase (HPO), which produces the antimicrobial hypohalous acid, has

previously been detected in ANG tissues, but at very low amounts, much less than the central core of the light organ, gills or ink sac, for comparison (29). Nitric oxide synthase (NOS) has been shown to be important in the establishment of the light organ and could play a role in the ANG. These reactive oxygen species, hypohalous acid and nitric oxide, may provide deterrents to keep nonsymbiotic bacteria out of the ANG and the symbionts are able to overcome these selective pressures.

Another potential antimicrobial peptides have also been shown to control bacterial symbionts, both in the guts of vertebrates and in the development of hydra (213, 214). No antimicrobial peptides have been thoroughly characterized in the squid, but they could provide an explanation for the profound partitioning of bacteria within the tubules of the ANG.

Despite characterizing the population of the ANG, its function is still unclear and is an exciting avenue for future research. Work by Allison Kerwin our lab has shown that eggs treated with antibiotics develop fungal biofilms (A. Kerwin, personal communication) suggesting that the bacteria from the ANG deposited in to the egg casings may produce antifungals. More research needs to be done, particularly to identify the putative antifungal compounds being produced. We have begun a collaboration with Dr. Marcy Balunas in order to extract and identify compounds produced by bacteria isolated from the ANG. Experiments must be done to disprove competing hypotheses, such as the bacteria on the egg casings may simply outcompete fungal and other fouling organisms for nutrients released from the eggs.

The last part of my doctoral work was to closely examine the bacterial species within the ANG, which was done primarily through comparative genomics. The most

readily culturable isolates from the ANG belong to the *Roseobacter* clade, an abundant and diverse lineage of marine *Alphaproteobacteria*. One species was repeatedly isolated from the ANGs of four different animals, *Phaeobacter sp.* ANG1. Colonies of *Phaeobacter sp.* ANG1 dominate plated dilutions of ANG homogenate. Moreover, the majority of 16S genes in sequenced clone libraries have the highest identity to this isolate, further supporting the conclusion that *Phaeobacter sp.* ANG1 is the dominant symbiont from the ANG. Future research should use this isolate as a model for the symbiosis, as it is the most readily culturable and dominates the ANG consortium. Gene manipulation is an invaluable tool in microbiology and it should be a priority to develop genetics in *Phaeobacter sp.* ANG1 as soon as possible.

To characterize some of the *Roseobacter* symbionts in the ANG, I used genome sequencing to compare the symbionts of the ANG to previously characterized organisms from the *Roseobacter* lineage. There were several interesting secondary metabolites produced by bacteria from the ANG consortium, including homoserine lactones and siderophores, which I further characterized. Homoserine lactones, being important quorum sensing molecules, are expressed both in culture and are detected in the ANG tissue. Future research must focus on the chemical characterization of these signaling molecules to confirm that they are of the same type produced in other organisms of the *Roseobacter* clade. More importantly, it must be determined what genes, if any, are controlled by these signaling molecules. Once genetics is available in *Phaeobacter sp.* ANG1 mutants lacking the autoinducer synthase and/or autoinducer receptor genes should be made. Given the previously characterized system that controls motility and biofilm formation in *Ruegeria sp.* KLH11 (43), experiments looking at these two

behaviors should be the first assays to be done. Transcriptomic analysis of the wild type, the quorum sensing deficient mutants may also shed light onto what genes are regulated by homoserine lactone production.

Many of the isolates from the ANG were found to contain siderophore biosynthetic genes. This is unusual for the *Roseobacter* lineage as very few other sequenced genomes contain siderophore synthesis genes. Testing the dominant symbiont, *Phaeobacter sp.* ANG1, against other siderophore producing roseobacters suggests that these symbionts have a growth advantage in iron limiting environments. ANG isolates readily produced siderophores as evidenced by measured CAS activity of the supernatants. Taken together, the ANG symbionts may outcompete competing members of the *Roseobacter* clade by secreting siderophores to acquire iron in the potentially iron-limiting environment of the ANG.

Much more work must be done to confirm this hypothesis. For example, while many *Roseobacter* clade organisms do not produce siderophores, many can transport siderophores and presumably acquire iron using siderophores produced by other organisms. The production of siderophores alone may not be a significant advantage in a mixed population of roseobacters. Furthermore, experiments must show that the ANG is actually an iron-limiting environment. Some iron chelating proteins are present in squid transcriptomic data, including ferritin. Experiments should determine if it and other proteins that sequester iron are present in the ANG. If the siderophore genes are expressed in the ANG, then it could be inferred that iron is limiting and siderophores are required for iron acquisition. Siderophores may also be only transiently expressed, particularly as the symbiosis is established.

Overall, results from my doctoral work open many new avenues for research in *Euprymna scolopes*. I have identified many new targets for research in the hemocytes of *E. scolopes* for further characterization. Characterizing the genes expressed in the primary immune cell of the squid is an important step to understand the molecular mechanisms of the innate immune system in the host. In addition, my work on the consortial symbiosis within the accessory nidamental gland opened many new and exciting avenues of research within the squid. At a time when symbiosis research is gaining widespread attention, particularly with animal-bacterial symbiosis, my doctoral work sets the stage for many future discoveries.

Appendix A. Draft Genome of *Maritalea sp.* RANG

Abstract

Maritalea sp. RANG is the first rhizobial isolate from the accessory nidamental gland of *Euprymna scolopes*. Presented here are features from a 3.4Mb assembly of its genome

Methods, Results, Discussion

The accessory nidamental gland (ANG) of cephalopods is dominated by *Alphaproteobacteria*. The ANG of *Euprymna scolopes* is no exception, with members of the *Phaeobacter* genus being the most abundant. While genomes of several of these have been sequenced, no genomes yet exist for the minority constituents of the ANG consortium, which include members of the *Rhizobiales*. The isolate *Maritalea sp.* RANG was isolated on a plate of Reasoner's 2A agar supplemented with sea salts and potassium nitrate and incubated anaerobically. The full-length 16S gene was sequenced and compared with 454 metagenomic 16S fragments which showed >98% identity to several of the reads, suggesting that this isolate is representative of the Rhizobia population within the ANG. When compared to characterized isolates, it had 97.6% identity to *Maritalea porphyrae*, originally isolated from red algae (215) and was given the preliminary designation as a member of the *Maritalea* genus.

The genome of *Maritalea sp.* RANG was sequenced by creating an Illumina Nextera-XT library and sequenced on a MiSeq Sequencer using 2 x 250 (Illumina, San Diego, CA). This resulted in 933, 244 reads which was assembled in to a 3.4 mB genome in 107 contigs using the CLC Genomic Workbench (CLC, Aarhus, Denmark) with approximately 68x coverage. 90% of the genome is contained in a single 3mB

contig. The assembled genome was submitted to the Rapid Annotation Sever using Subsystem Technology (RAST, 162) which identified 3,358 genes, including 38 structural RNA genes.

Maritalea sp. RANG has a complete pentose-phosphate pathway for carbohydrate metabolism and lacks the Embden-Meyerhof-Parnas and Entner-Doudoroff pathways. All enzymes for the TCA cycle are present. Contrary to the *Rhodobacterales* isolated from the ANG, *Maritalea sp.* RANG has enzymes to utilize amino sugars, including N-acetyl glucosamine, the monomer of chitin. While it has no chitinase, it may be able to metabolize subunits of chitin, if they are provided to the symbionts in the ANG.

Like many other ANG symbionts, *Maritalea sp.* RANG has the ability to respire nitrate anaerobically. Furthermore, it is able to carry out complete dentrification, reducing nitrate through several intermediates to nitrogen gas. It also is motile via a flagellum and has many genes related to chemotaxis, including 10 putative methyl accepting proteins (MCPs), responsible for sensing extracellular chemical gradients.

Maritalea sp. RANG also has many genes for dealing with potential stressors within the ANG. It has several enzyme to respond to reactive oxygen species, including catalase, superoxide dismutase, an alkyl hydroperoxidase subunit c-like protein (AhpC) and also genes involved in bacterial responses to nitric oxide, a key signaling molecule in the symbiosis between *Euprymna scolopes* and *V. fischeri* (27).

Being the first symbiont to be cultured from the accessory nidamental gland that is not from the *Roseobacter* lineage, *Maritalea sp.* RANG offers a unique perspective on a symbiosis that is conserved in many cephalopods.

Appendix B. Recombinant Expression of EsPGRP5

Introduction

Peptidoglycan recognition proteins (PGRPs) have been shown to be a critical pattern recognition receptor (PRR) in innate immunity. One of the defining characteristics of these proteins is the binding and/or degradation of peptidoglycan or its derivative. EsPGRP5 is a novel PGRP found in the hemocytes of *Euprymna scolopes*. While it has conserved residues that indicate amidase activity, this amidase activity must be proven biochemically. To accomplish this, two separate strategies were used to express recombinant EsPGRP5 (rEsPGRP5): baculovirus expression in *Spodoptera frugiperda* cells and lac-promoter driven expression in *Escherichia coli*. While both strategies succeeded in producing recombinant protein, purification from both of these methods was not successful. Further optimization of the purification protocols are needed to obtain the recombinant protein from crude cell lysate. Obtaining the recombinant protein is a critical step moving forward in hemocyte research, as EsPGRP5 is the most highly expressed PRR in circulating hemocytes.

Methods

Previous research suggested that expressing PGRPs in an *E. coli* expression system would be lethal to the cells and result in poor protein production (25). However, other research efforts have shown that both long-form and short-form PGRPs have successfully been expressed in *E. coli* (83, 216). Following on previous work done with PGRPs from *Euprymna scolopes*, EsPGRP5 was first expressed in the insect cell line SF9 from *Spodoptera frugiperda* using a modified version of the Bac-to-Bac expression platform (Life Technologies, Grand Island, NY).

Cloning of EsPGRP5 into baculovirus genome

RNA from adult gills were extracted using TRIzol (Life Technologies, Grand Island, NY) and reverse transcribed with the iScript cDNA synthesis kit (Bio-Rad, Hercules, CA). Primers were designed to amplify the 561 bp protein coding region of EsPGRP5 (Table 1). Restriction sites were added for directional cloning into the pFastBac vector containing transposon sites. The coding region of EsPGRP5 was amplified using Phusion 2X MasterMix (New England Biolabs, Ipswich, MA), which contains a proof-reading polymerase. Both pFastBac vector and the PCR product were treated with NcoI and XhoI then ligated together with T4 ligase. The ligation reaction was ethanol precipitated to remove salts and resuspended with 10 µL of water. Chemically competent DH5α cells were transformed with 2 µL of the ligation product and plated onto LB-ampicillin (100 µg/mL) agar plates. Transformants were analyzed by PCR.

A transformant containing the proper insertion was grown in a 5 mL overnight culture and plasmid was extracted with a Qiagen Miniprep kit (Hilden, Germany). The extracted plasmid was transformed into chemically-competent DH10Bac cells containing a baculovirus genome (or bacmid) with corresponding transposon recombination sites. Transformants were plated on LB agar plates containing kanamycin (50 µg/mL), tetracycline (10 µg/mL), gentamicin (7 µg/mL), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal, 100 µg/mL) and IPTG (40 µg/mL). The cloned gene on the pFastBac recombined with the bacmid, placing EsPGRP5 gene downstream of a viral-capsid promoter that is active after the baculovirus infects an insect cell. Inserting the gene into pFastBac also places a poly-histidine tag on the amino end of the protein and

disrupts β-galactosidase encoded on the plasmid. Resulting white transformants were screened by PCR and the correct sequence was confirmed by BigDye sequencing.

Successful transformants were grown in 100 mL cultures of LB with appropriate antibiotics and the bacmid was isolated by a modified version of the Qiagen Midi plasmid purification kit (Hilden, Germany).

Transfection of SF9 cells

Once purified bacmid was isolated, SF9 cells, derived from *Spodoptera frugiperda*, were transfected using 10μg of bacmid DNA combined with the lipid mixture CellFectin (Life Technologies, Grand Island, NY). SF9 cells were diluted to 5×10^5 and 2 mL of cells (1×10^6 in total) were added to 2 wells of a 6-well cell culture dish. Cells were allowed to adhere for 1 hour. Transfection mixture was added and plates were incubated for 5 h at 28°C to allow for infection. After this incubation, the transfection mixture was removed and 2 mL of growth media was added to the cells. Cells were incubated for 3 days at 28°C to allow generation of infectious viral particles. Supernatant containing virus was collected on the third day. The supernatant was centrifuged at 1,000 x g for 5 min to remove contaminating SF9 cells and then filtered through a 0.2 micron filter before being stored at 4°C.

To obtain higher viral titers, the harvested virus was amplified by a second round of infection. SF9 cells were grown to a density of 2×10^6 cells/mL. 50 mL of this culture was placed into a 1 L flask and infected with 1 mL of virus collected from the initial transfection. Cells were incubated at 28°C for 3 days with shaking at 150 rpm to maximize aeration. Supernatant was collected and purified of contaminants as described above.

Viral plaque assays

Viral plaque assays were performed on the amplified virus in order to determine proper infection of SF9 cells for protein production. $1x10^6$ SF9 cells were added to each well of a 6-well cell culture plate and allowed to adhere as before. Virus was serially diluted 10-fold and 100ul of one of the dilutions 10^{-4} to 10^{-8} was added to each well. The 6^{th} well was used as a negative control. Cells were incubated with virus for 1 hour at RT, with agitation every 15 min. The viral dilution was removed and replaced with 2mL of molten 1% agarose combined with SF-900 III complete media. Once the agarose had solidified, 1 mL of culture media was added to each well to prevent desiccation. Culture dishes were incubated for 4 days at 28°C. Plaques from each dilution were counted and the undiluted cell titer was determined. In my hands, good virus amplification yielded greater than 1×10^8 plaque forming units/mL.

Protein expression

Large scale protein production was done by growing 500 mL of SF9 cells to a density of 2×10^6 cells/mL and infecting with virus with a multiplicity of infection (MOI, equal to the number of viral particles per cell) of ~2, which was approximately 20 mL amplified virus stock. Cells were incubated for 3 days in a 2 L flask at 28°C with shaking at 150rpm. Cells were pelleted at 1,000 x g for 5 min and frozen at -80°C until needed

Purification of recombinant protein

Cells were lysed with lysis buffer (1% TRITON, 20 mM NaPO4, 0.5 M NaCl, 20 mM imidazole) and centrifuged at 20,000 x g for 30 min. The supernatant was removed and the centrifugation was repeated. The resulting supernatant was filtered through a

0.45 micron SFCA filter to remove impurities. The histidine-tagged protein was purified on a 1 mL nickel affinity column (GE Healthcare, Little Chalfont, UK). Fractions were collected at all steps for further analysis. The column was equilibrated with 5 volumes of binding buffer (20 mM NaPO4, 0.5 M NaCl, 20 mM imidazole), the cell lysate was applied to the column and washed with 10 volumes of binding buffer. The protein was eluted with 5 volumes of elution buffer (20 mM NaPO4, 0.5 M NaCl, 0.5 M imidazole).

Recombinant protein expression in *E. coli*

EsPGRP5 was cloned from cDNA synthesized from hemocyte RNA as described above. The restriction site on the forward primer was altered to make it compatible with the pET28a expression vector (Table 1). The PCR product and the vector were treated with restriction enzymes NdeI and XhoI and ligated into the expression vector pET28a with T4 ligase. Ligation reactions were ethanol precipitated to remove salt and resuspended in 10 µL of water. BL21 *E. coli* were transformed with 2 µL of this ligation via electroporation. Transformants were selected for on LB-kanamycin (50µg/mL) agar plates and insertion was confirmed using PCR and sequencing using BigDye sequencing reagent (Life Technologies, Grand Island, NY).

A successful transformant was isolated and protein expression was induced in one of two ways. First, after an overnight culture was grown in LB, an aliquot was used to inoculate LB media and allowed to grow to an OD_{600} of 0.6 before inducing with 1mM of IPTG. Cells were collected after 4 hours of incubation by centrifugation at 10,000 x g. Second, an overnight culture of the transformant grown in LB was used to inoculate autoinducing medium as described previously (217) which takes advantage of catabolite repression to induce protein expression. These induced cells were grown overnight at

either 37°C, 20°C or 25°C before the cells were collected by centrifugation at 10,000 x g. The supernatant was removed and the cell pellet frozen at -20°C.

Protein purification from *E. coli*

Frozen cell pellets were resuspended in 20 mL of lysis buffer (50mM Na_2PO_4, 50 mM NaCl, pH = 8), then thawed and treated with lysozyme at 0.2 mg/ml for 2 hours at room temperature. Cells were chilled on ice and sonicated with a Branson Sonifier Cell Disrupter 200 on power level 4 for 20 second bursts. Sonication was repeated, keeping cells on ice for 1 minute intervals between bursts, until cells were visibly lysed when examined under a microscope. Aliquots of cell lysate were saved as "total protein" samples.

Cell lysate was clarified by centrifugation at 30,000 x g for 20 min. The resulting supernatant was filtered through a 0.45 micron SFCA filter. Aliquots of filtered supernatant were saved as "soluble protein" samples.

Analysis of protein purification

To analyze the protein present in any given fraction from protein purification, 15 µL aliquots of sample was mixed with Laemmli buffer and heated at 95°C for 5 min. These samples were run on a 12% TGX polyacrylamide gel (Bio-Rad) at 200V for 35 min. Gels were stained overnight in Coomassie R-250 stain (1 g of R-250 dye in 500 mL methanol, 100 mL acetic acid and 400 mL water) and destained the following day with destain (R-250 staining solution without the R-250 dye added). Some gels were also fixed in a solution of 30% EtOH and 2% phosphoric acid overnight and the next day stained for 4-6 hours with G-250 protein stain (10% EtOH, 2% phosphoric acid, 0.25% Coomassie G-250). If needed, these gels were destained in gel fixative to eliminate

background staining. For increased sensitivity, gels were stained with SYPRO Ruby (Bio-Rad, Hercules, CA) and imaged with UV illumination on a Bio-Rad Gel Doc. Western blots were done as described previously with hemocyte protein extracts (See Chapter 2).

Results and Discussion

Expression of protein in either system, baculovirus or IPTG/lactose-induced *E. coli*, proved capable of producing recombinant EsPGRP5. Yield from SF9 cells seemed to be lower in terms of total protein. However, a clear band was seen at 25kD, which is the approximate size of the EsPGRP5 protein (21.2kD) plus the added polyhistidine tag and tobacco etch virus (TEV) cleavage site (Figure 1). A Western blot then confirmed that the recombinant protein was present in the cell lysate (Figure 2, lane 1). However, after the sample was put through a nickel column, no protein was detected by Coomassie stain. Western blot detected the recombinant protein in the flow through and eluted fractions, suggesting that very little of the recombinant protein bound to the purification column (Figure 2). Furthermore, a SYPRO Ruby stain of an elution fraction showed a general banding pattern in the elution fractions, showing that many proteins had some affinity to the column and there was no selection for the polyhistidine tag (Figure 3).

By comparison, recombinant protein expression in *E. coli* produced a significant amount of protein that was visible by Coomassie staining (Figure 4). Unfortunately, the majority of protein seemed to be insoluble, suggesting the expressed protein was present in inclusion bodies. Efforts to correct this with autoinducing media and incubation at different temperatures were not successful in increasing the amount of soluble recombinant protein (Figure 4). Attempts to purify the small amount of soluble protein in

these samples were not successful. Coomassie staining of the elution fractions showed non-specific binding of protein (Figure 5). If more soluble recombinant protein had been solubilized, the polyhistidine tags on the recombinant protein may have outcompeted the non-specific protein binding to the nickel column and the yield of purified protein may have been greater.

Given the significant amount of effort required to express protein in SF9 cells without dedicated incubators, future work should continue to express rPGRP5 in *E. coli*. Similar proteins have been expressed by cloning the gene into an expression vector that is induced by IPTG, such as pET28a, then solubilizing the protein with urea. This will allow the protein to bind to the nickel column and once refolded, will be a functional protein for analysis. This work should be done for EsPGRP5 as it is one of the most highly expressed genes in circulating hemocytes. Its role in innate immunity must be elucidated and biochemical assays would do the most to ascertain the function of this protein within *E. scolopes*.

Table 1. Primers used in recombinant protein expression of EsPGRP5

Name	Sequence (5’ to 3’)
PGRP5S-F	CAGTCCATGGATAGGTGGAGGAAATC
PGRP5R	CAGTCTCGAGTTATGTAATGGGGCCA
PGRP5pET28F	CAGTCATATGATGCAAGGGTGGAGGAAATC

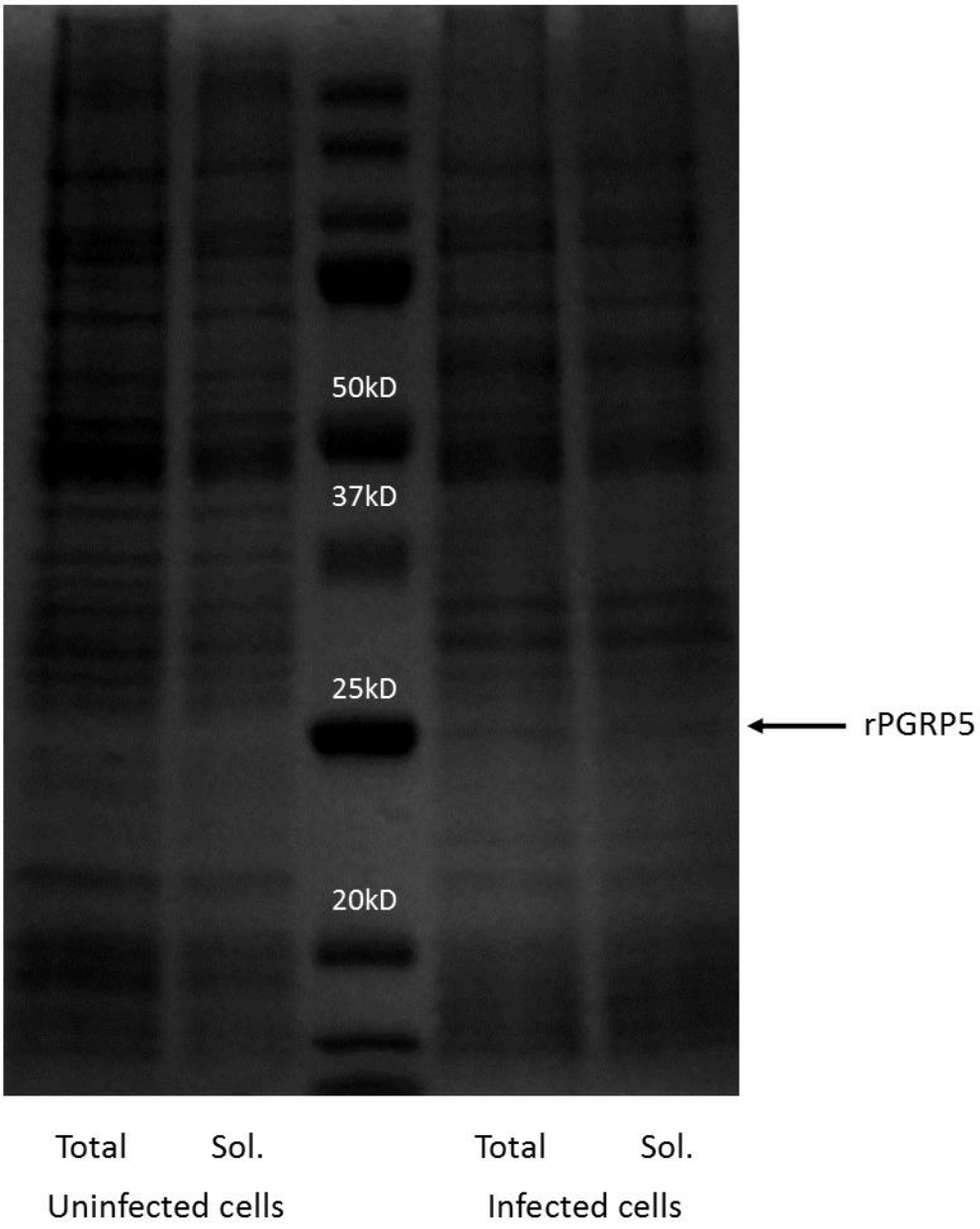

Figure 1. Protein expression of rPGRP5 in SF9 cells using baculovirus. A 25kD band appeared in protein samples from infected cells, indicative of expression of the 21.2.kD EsPGRP5 protein with the polyhistidine tag, tobacco etch virus (TEV) protease cleavage site and linker peptide

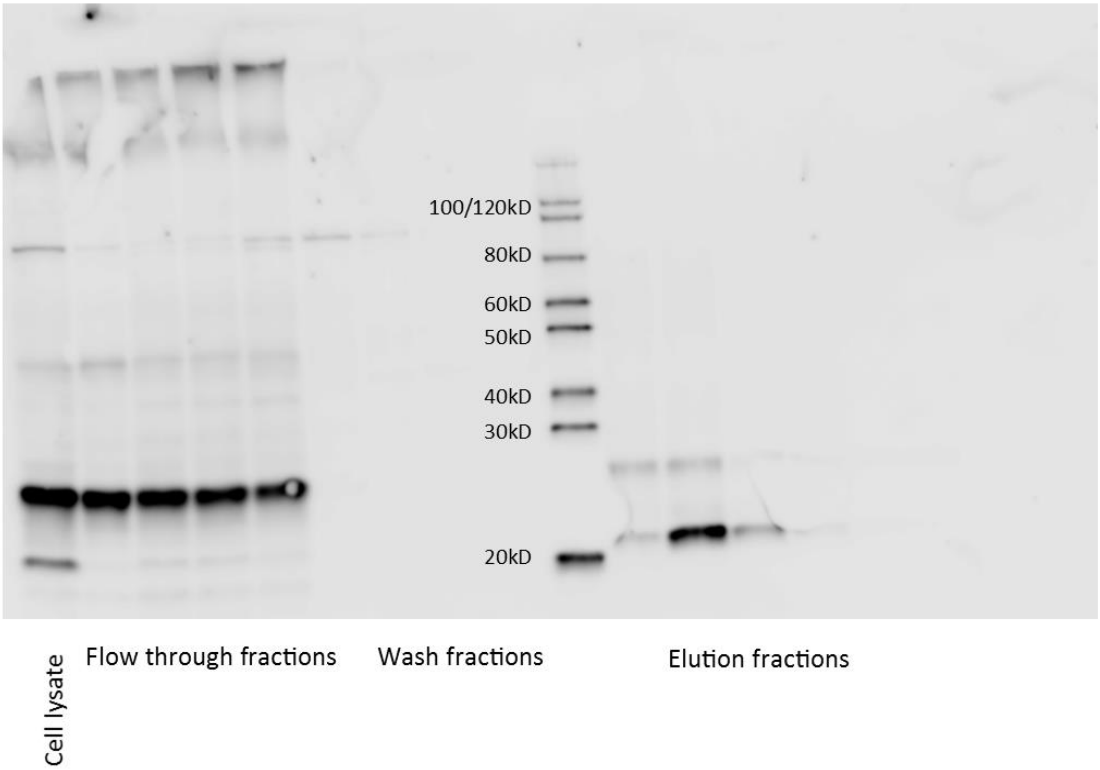

Figure 2. Western blot of SF9 protein fractions shows inefficient purification of recombinant protein. While a large amount of protein was detected in cell lysate (Lane 1) The majority of recombinant protein did not bind to the nickel column and was detected in the flow through (lanes 2-5). Two distinct bands were detected in the elution fractions (Lanes 11-15); a 25kD protein corresponding to the full length recombinant rPGRP5 protein and a 22kD, possibly corresponding to a recombinant protein with the polyhistidine tag cleaved off.

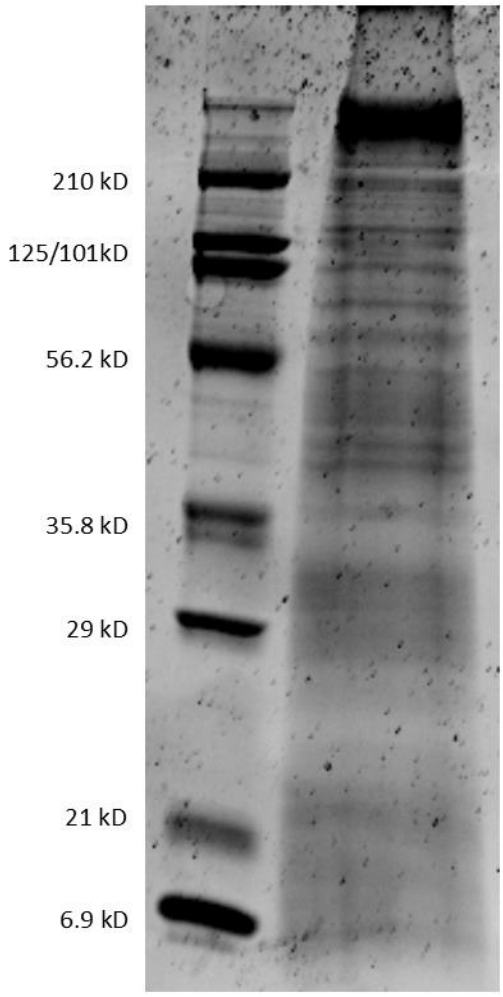

Figure 3. SYPRO Ruby staining of nickel column elution fraction shows no selection for polyhistidine-tagged protein. A sample of elution fraction #2 (Lane 12 in Figure 2) analyzed by SDS-PAGE and stained with SYPRO Ruby fluorescent stain, with nearly the sensitivity of a silver stain. The high number of protein bands and smearing patterns suggest that the protein purification protocol is not effective at selecting for polyhistidine-tagged proteins. Protein degradation may also be a factor.

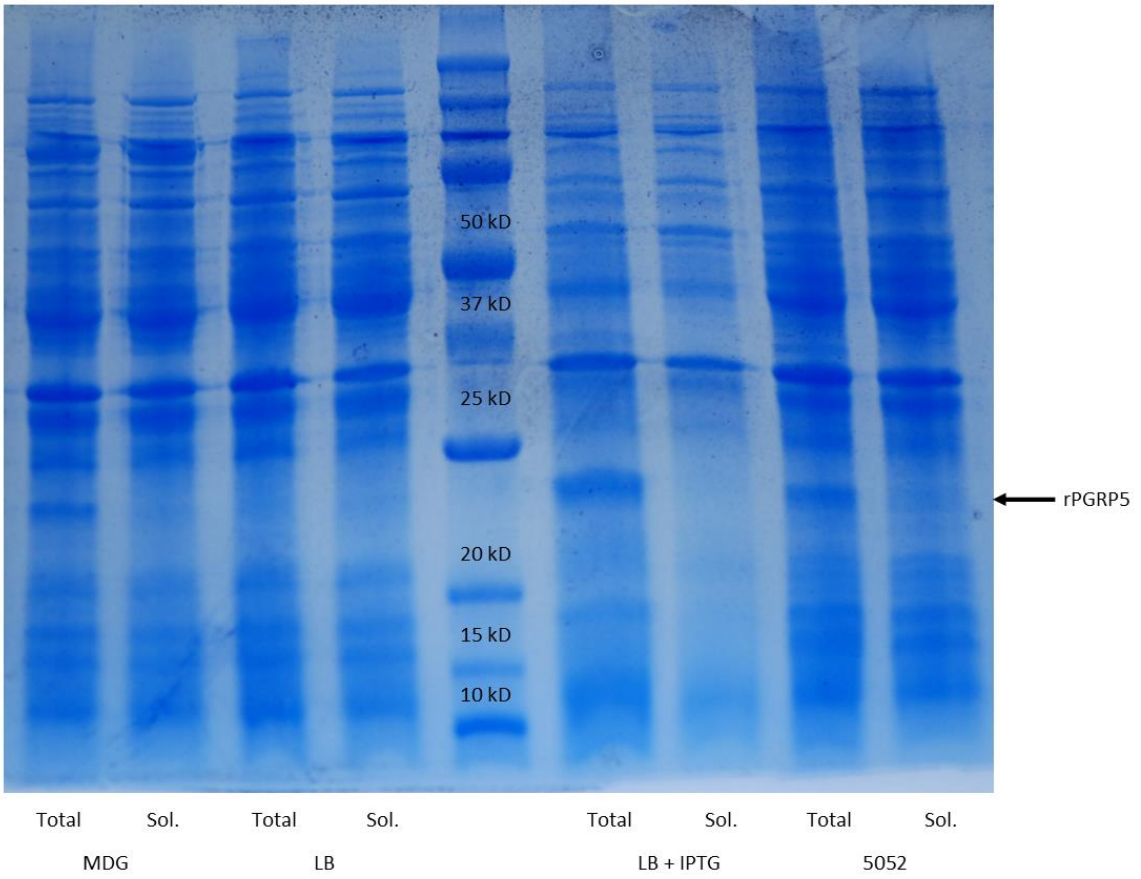

Figure 4. EsPGRP5 expression in *E. coli* using different induction strategies. *E. coli* BL21 carrying EsPGRP5 in the expression plasmid pET28a was grown and induced under several different conditions. Protein expression was induced either with IPTG (LB+IPTG) or with lactose (5052 media). Two different negative controls were used that contained neither lactose nor IPTG. A band in the MDG control suggests lactose contamination of the media. In all cases, the majority of rEsPGRP5 appeared in total protein samples, suggesting that rEsPGRP5 was accumulating in insoluble inclusion bodies.

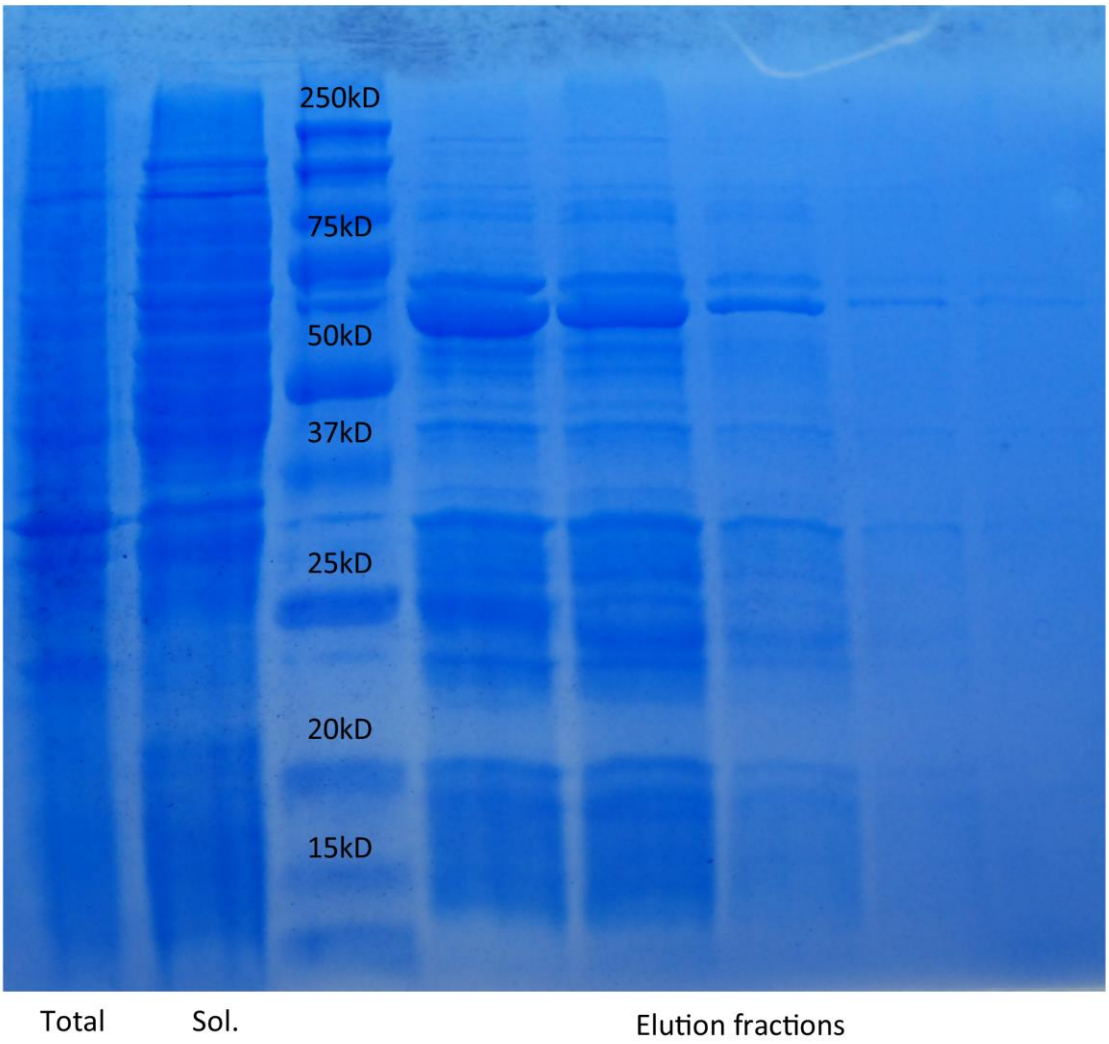

Figure 5. EsPGRP5 is not recovered from *E. coli* lysate. Despite an abundance of rEsPGRP5 in the total protein fraction of *E. coli* lysate, very little was seen in the soluble fraction (Sol.). Attempts to purify the little recombinant protein from the soluble fraction was not effective, as non-specifically bound proteins predominated the elution fractions.

References

1. **Whitman WB, Coleman DC, Wiebe WJ**. 1998. Prokaryotes: the unseen majority. Proc. Nat. Acad. Sci. U.S.A. **95:**6578-83.

2. **Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC.** 2007. How rhizobial symbionts invade plants: the *Sinorhizobium-Medicago* model. Nat. Rev. Microbiol. **5:** 619-33.

3. **Sellstedt A, Richau KH.** 2013. Aspects of nitrogen-fixing actinobacteria, in particular free-living and symbiotic. FEMS Microbiol. Lett. **342**:179-186.

4. **Walter J, Ley R.** 2011. The human gut microbiome: ecology and recent evolutionary changes. Annu. Rev. Microbiol. **65**:411-29.

5. **Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI.** 2005. Host-bacterial mutualism in the human intestine. Science **307**:1915-20.

6. **Bry L, Falk PG, Midtvedt T, Gordon JI.** 1996. A model of host-microbial interactions in an open mammalian ecosystem. Science **273**:1380-3.

7. **Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, Srinivasan S, Sitaraman SV, Knight R, Ley RE, Gewirtz AT.** 2010. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor. Science **328**:228-31.

8. **Bates JM, Mittge E, Kuhlman J, Baden KN, Cheesman SE, Guillemin K.** 2006. Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. Dev. Bio. **297**:374-386.

9. **Nelson MC, Graf J.** 2012. Bacterial symbioses of the medicinal leech *Hirudo verbana*. Gut Microbes **3**:322-331.

10. **Graf J, Kikuchi Y, Rio RVM.** 2006. Leeches and their microbiota: naturally simple symbiosis models. Trends Microbiol. **14**:365-371.

11. **McFall-Ngai MJ, Ruby EG.** 1991. Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. Science **254**:1491-4.

12. **Ruby EG, McFall-Ngai MJ.** 1992. A squid that glows in the night: development of an animal-bacterial mutualism. J. Bacteriol. **174**:4865-70.

13. **Bomar L, Maltz M, Colston S, Graf J.** 2011. Directed culturing of microorganisms using metatranscriptomics. MBio **2**:e00012-11-e00012-11.

14. **Silver AC, Kikuchi Y, Fadl AA, Sha J, Chopra AK, Graf J.** 2007. Interaction between innate immune cells and a bacterial type III secretion system in mutualistic and pathogenic associations. Proc. Nat. Acad. Sci. U.S.A **104**:9481-9486.

15. **Kikuchi Y, Graf J.** 2007. Spatial and temporal population dynamics of a naturally occurring two-species microbial community inside the digestive tract of the medicinal leech. Appl. Environ. Microbiol. **73**:1984-91.

16. **Jones B, Nishiguchi M**. 2004. Counterillumination in the hawaiian bobtail squid, *Euprymna scolopes* Berry (Mollusca: Cephalopoda). Mar. Biol. **144**:1151-1155

17. **Foster JS, Apicella MA, McFall-Ngai MJ**. 2000. *Vibrio fischeri* lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. Dev. Biol. **226**:242-254.

18. **Koropatnick TA, Engle JT, Apicella MA, Stabb EV, Goldman WE, McFall-Ngai MJ.** 2004. Microbial factor-mediated development in a host-bacterial mutualism. Science **306**:1186-8.

19. **Doino J, McFall-Ngai M.** 1995. A transient exposure to symbiosis-competent bacteria induces light organ morphogenesis in the host squid. Biol. Bull. **187**: 347-355.

20. **Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ.** 2002. Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. Appl. Environ. Microbiol. **68**:5113-5122.

21. **Nyholm SV, McFall-Ngai M.** 2004. The winnowing: establishing the squid–vibrio symbiosis. Nat. Rev. Microbiol. **2**:632-642.

22. **Lupp C, Ruby EG.** 2005. *Vibrio fischeri* uses two quorum-sensing systems for the regulation of early and late colonization factors. J. Bacteriol. **187**:3620-3629.

23. **Nyholm S, McFall-Ngai M.** 1998. Sampling the light-organ microenvironment of Euprymna scolopes: description of a population of host cells in association with the bacterial symbiont *Vibrio fischeri*. Biol. Bull. **195**: 89-97.

24. **Wier AM, Nyholm SV, Mandel MJ, Massengo-Tiasse RP, Schaefer AL, Koroleva I, Splinter-BonDurant S, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, Bonaldo MD, Casavant TL, Soares MB, Cronan JE, Reed JL, Ruby EG, McFall-Ngai MJ.** 2010. Transcriptional patterns in both host and bacterium underlie a daily rhythm of anatomical and metabolic change in a beneficial symbiosis. Proc. Nat. Acad. Sci. U.S.A. **107**:2259-2264.

25. **Troll JV, Bent EH, Pacquette N, Wier AM, Goldman WE, Silverman N, McFall-Ngai MJ.** 2010. Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. Environ. Microbiol. **12**:2190-203.

26. **Xu W, Banchereau J.** 2014. The antigen presenting cells instruct plasma cell differentiation. Front. Immunol. **4**:10.3389/fimmu.2013.00504

27. **Davidson SK, Koropatnick TA, Kossmehl R, Sycuro L, McFall-Ngai MJ.** 2004. NO means “yes” in the squid-vibrio symbiosis: nitric oxide (NO) during the initial stages of a beneficial association. Cell. Microbiol. **6**:1139-51.

28. **Weis VM, Small AL, McFall-Ngai MJ.** 1996. A peroxidase related to the mammalian antimicrobial protein myeloperoxidase in the *Euprymna-Vibrio* mutualism. Proc. Natl. Acad. Sci. U.S.A. **93**:13683-8.

29. **Small AL, McFall-Ngai MJ.** 1999. Halide peroxidase in tissues that interact with bacteria in the host squid *Euprymna scolopes*. J. Cell. Biochem. **72**:445-57.

30. **Visick KL, Ruby EG.** 1998. The periplasmic, group III catalase of *Vibrio fischeri* is required for normal symbiotic competence and is induced both by oxidative stress and by approach to stationary phase. J. Bacteriol. **180**:2087-92.

31. **Ganz T.** 2009. Iron in innate immunity: starve the invaders. Curr. Opin. Immunol. **21**:63-67.

32. **Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Rol, Strong RK, Strong K, Akira S, Aderem A.** 2004. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. Nature **432**:917-21.

33. **Koropatnick T, Kimbell J, McFall-Ngai M.** 2007. Responses of host hemocytes during the initiation of the squid-*Vibrio* symbiosis. Biol. Bull. **212**:29-39.

34. **Nyholm SV, Stewart JJ, Ruby EG, McFall-Ngai MJ.** 2009. Recognition between symbiotic *Vibrio fischeri* and the haemocytes of *Euprymna scolopes*. Environ. Microbiol. **11**:483-93.

35. **Wagner-Döbler I, Ballhausen B, Berger M, Brinkhoff T, Buchholz I, Bunk B, Cypionka H, Daniel R, Drepper T, Gerdts G, Hahnke S, Han C, Jahn D, Kalhoefer D, Kiss H, Klenk H, Kyrpides N, Liebl W, Liesegang H, Meincke L, Pati A, Petersen J, Piekarski T, Pommerenke C, Pradella S, Pukall R, Rabus R, Stackebrandt E, Thole S, Thompson L, Tielen P, Tomasch J, Jan von M, Wanphrut N, Wichels A, Zech H, Simon M.** 2010. The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. ISME J **4**:61-77.

36. **Shiba T.** 1991. *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll a. Syst. Appl. Microbiol. **14:** 140-145.

37. **Kiene RP, Linn LJ.** 2000. Distribution and turnover of dissolved DMSP and its relationship with bacterial production and dimethylsulfide in the Gulf of Mexico. 4. Limnol.Oceanogr. **45**:849-861.

38. **Malmstrom RR, Kiene RP, Kirchman DL.** 2004. Identification and enumeration of bacteria assimilating dimethylsulfoniopropionate (DMSP) in the north Atlantic and Gulf of Mexico. Limnol. Oceanogr. **49**:597-606.

39. **Charlson R, Lovelock J, Andreae M, Warren S.** 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. Nature. **326**:655-661.

40. **Gonzalez JM, Simo R, Massana R, Covert JS, Casamayor EO, Pedros-Alio C, Moran MA.** 2000. Bacterial Community Structure Associated with a Dimethylsulfoniopropionate-Producing North Atlantic Algal Bloom. Appl. Environ. Microbiol. **66**:4237-4246.

41. **Vila-Costa M, del Valle DA, González JM, Slezak D, Kiene RP, Sánchez O, Simó R.** 2006. Phylogenetic identification and metabolism of marine dimethylsulfide-consuming bacteria. Environ. Microbiol. **8**:2189-2200.

42. **Dang H, Li T, Chen M, Huang G.** 2008. Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. Appl. Environ. Microbiol. **74**:52-60.

43. **Zan J, Cicirelli EM, Mohamed NM, Sibhatu H, Kroll S, Choi O, Uhlson CL, Wysoczynski CL, Murphy RC, Churchill MEA, Hill RT, Fuqua C.** 2012. A complex LuxR-LuxI type quorum sensing network in a roseobacterial marine sponge symbiont activates flagellar motility and inhibits biofilm formation. Mol. Microbiol. **85**:916-933.

44. **Rao D, Webb JS, Holmström C, Case R, Low A, Steinberg P, Kjelleberg S.** 2007. Low densities of epiphytic bacteria from the marine alga *Ulva australis* inhibit settlement of fouling organisms. Appl. Environ. Microbiol. **73**:7844-52.

45. **Berger M, Neumann A, Schulz S, Simon M, Brinkhoff T.** 2011. Tropodithietic acid production in *Phaeobacter gallaeciensis* is regulated by N-acyl homoserine lactone-mediated quorum sensing. J. Bacteriol. **193**:6576-85.

46. **D'Alvise PW, Lillebø S, Prol-Garcia MJ, Wergeland HI, Nielsen KF, Bergh Ø, Gram L.** 2012. *Phaeobacter gallaeciensis* reduces *Vibrio anguillarum* in cultures of microalgae and rotifers, and prevents vibriosis in cod larvae. PLoS ONE **7**:e43996.

47. **Ruiz-Ponte C, Cilia V, Lambert C, Nicolas JL.** 1998. Roseobacter gallaeciensis sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus.* Int. J. Syst. Bacteriol. **2**:537-42.

48. **Grigioni S, Boucher-Rodoni R, Demarta A, Tonolla M.** 2000. Phylogenetic characterisation of bacterial symbionts in the accessory nidamental glands of the sepioid *Sepia officinalis* (Cephalopoda: Decapoda). Mar. Biol. **136**:217-222.

49. **Barbieri E, Paster BJ, Hughes D, Zurek L, Moser DP, Teske A, Sogin ML.** 2001. Phylogenetic characterization of epibiotic bacteria in the accessory nidamental gland and egg capsules of the squid *Loligo pealei* (Cephalopoda:Loliginidae). Environ. Microbiol. **3**:151-67.

50. **Collins AJ, LaBarre BA, Won BSW, Shah MV, Heng S, Choudhury MH, Haydar SA, Santiago J, Nyholm SV.** 2012. Diversity and partitioning of bacterial populations within the accessory nidamental gland of the squid *Euprymna scolopes*. Appl. Environ. Microbiol. **78**:4200-8.

51. **Pichon D, Gaia V, Norman M, Boucher-Rodoni R.** 2005. Phylogenetic diversity of epibiotic bacteria in the accessory nidamental glands of squids (Cephalopoda: Loliginidae and Idiosepiidae). Mar. Biol. **147**:1323:1332.

52. **Sagan L.** 1967. On the origin of mitosing cells. J. Theor. Biol. **14**:255-74.

53. **Gross R, Vavre F, Heddi A, Hurst GD, Zchori-Fein E, Bourtzis K.** 2009. Immunity and symbiosis. Mol. Microbial. **73**:751-759.

54. **McFall-Ngai M, Nyholm SV, Castillo MG.** 2010. The role of the immune system in the initiation and persistence of the *Euprymna scolopes--Vibrio fischeri* symbiosis. Semin. Immunol. **22**:48-53.

55. **Royet J, Gupta D, Dziarski R.** 2011. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. Nat. Rev. Immunol. **11**:837-51.

56. **McFall-Ngai M.** 2007. Adaptive immunity: care for the community. Nature. **445**:153.

57. **Lee YK, Mazmanian SK.** 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? Science **330**:1768-1773.

58. **Bauer E, Williams BA, Smidt H, Verstegen MWA, Mosenthin R.** 2006. Influence of the gastrointestinal microbiota on development of the immune system in young animals. Curr. Issues Intest. Microbiol. **7**:35-51.

59. **Olszak T, An D, Zeissig S, Vera MP, Richter J, Franke A, Glickman JN, Siebert R, Baron RM, Kasper DL, Blumberg RS.** 2012. Microbial exposure during early life has persistent effects on natural killer T cell function. Science **336**:489-493.

60. **Duan J, Kasper DL.** 2011. Regulation of T cells by gut commensal microbiota. Curr. Opin. Rheumatol. **23**:372-376.

61. **Stewart FJ, Cavanaugh CM.** 2006. Symbiosis of thioautotrophic bacteria with *Riftia pachyptila.* Prog. Mol. Subcell. Biol. **41**:197-225.

62. **Moran NA, McCutcheon JP, Nakabachi A.** 2008. Genomics and evolution of heritable bacterial symbionts. Annu. Rev. Genet. **42**:165-190.

63. **Fedlhaar H, Gross R**. 2009. Insects as hosts for mutualistic bacteria. Int. J. Med. Microbiol. **299**:1-8.

64. **Watanabe H, Tokuda G.** 2010. Cellulolytic systems in insects. Ann. Rev. Entomol. **55**:609-632.

65. **Backhed F, Manchester JK, Semenkovich CF, Gordon JI.** 2007. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. Proc. Nat. Acad. Sci. U.S.A **104**:979-984.

66. **Stoll S, Feldhaar H, Fraunholz MJ, Gross R.** 2010. Bacteriocyte dynamics during development of a holometabolous insect, the carpenter ant *Camponotus floridanus*. BMC Microbiol. **10**:308.

67. **Krasity B, Troll J, Weiss J, McFall-Ngai MJ**. 2011. LBP/BPI proteins and their relatives: conservation over evolution and roles in mutualism. Biochem. Soc. Trans. **39**:1039-1044.

68. **Weiss BL, Wang J, Aksoy S**. 2011. *Tsetse* immune system maturation requires the presence of obligate symbionts in larvae. PLoS Biol. **9**:e1000619.

69. **Visick KL, Ruby EG.** 2006. *Vibrio fischeri* and its host: it takes two to tango. Curr. Opin. Microbiol. **9**:632-638.

70. **McFall-Ngai, MJ.** 2002. Unseen forces: the influence of bacteria on animal development. Dev. Biol. **242**:1-14.

71. **Castillo MG, Goodson MS, McFall-Ngai MJ.** 2009. Identification and molecular characterization of a complement C3 molecule in a lophotrochozoan, the Hawaiian bobtail squid *Euprymna scolopes*. Dev. Comp. Immunol. **33**:69-76.

72. **Troll JV, Adin DM, Wier AM, Paquette N, Silverman N, Goldman WE, Stadermann FJ, Stabb EV, McFall-Ngai MJ.** 2009. Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. Cell. Microbiol. **11**:1114-27.

73. **McFall-Ngai M, Heath-Heckman EA, Gillette AA, Peyer SM, Harvie EA.** 2012. The secret languages of coevolved symbioses: Insights from the *Euprymna scolopes*–*Vibrio fischeri* symbiosis. Semin. Immunol. **24**:3-8.

74. **Heath-Heckman EAC, McFall-Ngai MJ.** 2011. The occurrence of chitin in the hemocytes of invertebrates. Zoology **114**:191-198.

75. **Schleicher TR, Nyholm SV.** 2011. Characterizing the host and symbiont proteomes in the association between the Bobtail squid, *Euprymna scolopes*, and the bacterium, *Vibrio fischeri*. PLoS ONE **6**:e25649.

76. **Collins A, Nyholm S.** 2010. Obtaining hemocytes from the Hawaiian bobtail squid *Euprymna scolopes* and observing their adherence to symbiotic and non-symbiotic bacteria. J. Vis. Exp. **11**: 10.3791/1714

77. **Lamarcq LH, McFall-Ngai MJ.** 1998. Induction of a gradual, reversible morphogenesis of its host's epithelial brush border by *Vibrio fischeri*. Infect. Immun. **66**:777-85.

78. **Petersen TN, Brunak S, Heijne G, Nielsen H, Heijne von G.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods **8**:785-786.

79. **Zdobnov EM, Apweiler R**. 2001. InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17**:847-8.

80. **Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ.** 1998. Multiple sequence alignment with Clustal X. Trends Biochem. Sci. **23**:403-5.

81. **Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M.** 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21**:3674-3676.

82. **Mellroth P.** 2003. A scavenger function for a drosophila peptidoglycan recognition protein. J. Biol. Chemist. **278**:7059-7064.

83. **Kim M, Byun M, Oh B**. 2003. Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster*. Nat. Immunol. **4**:787-793.

84. **Vasta GR.** 2009. Roles of galectins in infection. Nat. Rev. Microbiol. **7**:424-38.

85. **Tasumi S, Vasta GR**. 2007. A galectin of unique domain organization from hemocytes of the eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite *Perkinsus marinus*. J. Immunol. **179**:3086-3098.

86. **Song X, Zhang H, Wang L, Zhao J, Mu C, Song L, Qiu L, Liu X.** 2011. A galectin with quadruple-domain from bay scallop Argopecten irradians is involved in innate immune response. Dev. Comp. Immunol. **35**:592-602.

87. **Zhang D, Jiang S, Hu Y, Cui S, Guo H, Wu K, Li Y, Su T**. 2011. A multidomain galectin involved in innate immune response of pearl oyster *Pinctada fucata*. Dev. Comp. Immunol. **35**:1-6.

88. **Medzhitov R**. 2007. Recognition of microorganisms and activation of the immune response. Nature **449**:819-826.

89. **Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB, McFall-Ngai MJ.** 2005. Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. Appl. Environ. Microbiol. **71**:6934-46.

90. **Leulier F, Lemaitre B.** 2008. Toll-like receptors — taking an evolutionary approach. Nat. Rev. Genet. **9**:165-178.

91. **Chun CK, Scheetz TE, de Bonaldo MF, Brown B, Clemens A, Crookes-Goodson WJ, Crouch K, DeMartini T, Eyestone M, Goodson MS, Janssens B, Kimbell JL, Koropatnick TA, Kucaba T, Smith C, Stewart JJ, Tong D, Troll JV, Webster S, Winhall-Rice J, Yap C, Casavant TL, McFall-Ngai MJ, Soares MB.** 2006. An annotated cDNA library of juvenile *Euprymna scolopes* with and without colonization by the symbiont Vibrio fischeri. BMC Genomics **7**:154.

92. **Kenny E, O'Neill L.** 2008. Signaling adaptors used by Toll-like receptors: An update. Cytokine **43**:342-349.

93. **Wells JM, Rossi O, Meijerink M, van Baarlen P.** 2011. Epithelial crosstalk at the microbiota-mucosal interface. Proc.Nat. Acad. Sci. U.S.A. **108**:4607-4614.

94. **Marinis JM, Homer CR, McDonald C, Abbott DW.** 2011. A novel motif in the crohn's disease susceptibility protein, NOD2, allows TRAF4 to down-regulate innate immune responses. J. Biol. Chemist. **286**:1938-1950.

95. **West AP, Brodsky IE, Rahner C, Woo DK, Erdjument-Bromage H, Tempst P, Walsh MC, Choi Y, Shadel GS, Ghosh S.** 2011. TLR signalling augments macrophage bactericidal activity through mitochondrial ROS. Nature 472:476-480.

96. **Conboy IM, Manoli D, Mhaiskar V, Jones PP.** 1999. Calcineurin and vacuolar-type H+-ATPase modulate macrophage effector functions. Proc. Natl. Acad. Sci. U.S.A. **96**:6324-9.

97. **Jennings C, Kusler B, Jones PP.** 2009. Calcineurin inactivation leads to decreased responsiveness to LPS in macrophages and dendritic cells and protects against LPS-induced toxicity in vivo. Innate Immun. **15**:109-120.

98. **de Lorgeril J, Zenagui R, Rosa RD, Piquemal D, Bachère E.** 2011. Whole transcriptome profiling of successful immune response to *Vibrio* infections in the oyster *Crassostrea gigas* by digital gene expression analysis. PLoS ONE **6**:e23142.

99. **Philipp EER, Kraemer L, Melzner F, Poustka AJ, Thieme S, Findeisen U, Schreiber S, Rosenstiel P.** 2012. Massively parallel RNA sequencing identifies a complex immune gene repertoire in the lophotrochozoan *Mytilus edulis*. PLoS ONE **7**:e33091.

100. **Dunkelberger JR, Song W.** 2009. Complement and its role in innate and adaptive immune responses. Cell Res. **20**:34-50.

101. **Aoun RB, Hetru C, Troxler L, Doucet D, Ferrandon D, Matt N.** 2011. Analysis of thioester-containing proteins during the innate immune response of *Drosophila melanogaster*. J. Innate Immun. 3:52-64.

102. **Blandin S.** 2004. Thioester-containing proteins and insect immunity. Mol. Immunol. **40**:903-908.

103. **Sottrup-Jensen L**. 1989. Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. J. Biol. Chemist. **264**:11539-42.

104. **Schvartz I, Seger D, Shaltiel S.** 1999. Vitronectin. Int. J. Biochem. Cell Biol. **31**:539-44.

105. **Singh B, Su Y, Riesbeck K.** 2010. Vitronectin in bacterial pathogenesis: a host protein used in complement escape and cellular invasion. Mol. Microbial. **78**:545-560.

106. **Ludin P, Nilsson D, Mäser P.** 2011. Genome-wide identification of molecular mimicry candidates in parasites. PLoS ONE **6**:e17546.

107. **Ueda A, Nagai H, Ishida M, Nagashima Y, Shiomi K.** 2008. Purification and molecular cloning of SE-cephalotoxin, a novel proteinaceous toxin from the posterior salivary gland of cuttlefish *Sepia esculenta*. Toxicon **52**:574-81.

108. **Kemper C, Atkinson JP, Hourcade DE.** 2010. Properdin: Emerging Roles of a Pattern-Recognition Molecule. Ann. Rev. Immunol. **28**:131-155.

109. **Kemper C, Hourcade DE.** 2008. Properdin: New roles in pattern recognition and target clearance. Mol. Immunol. **45**:4048-4056.

110. **Shiomi K, Midorikawa S, Ishida M, Nagashima Y, Nagai H.** 2004. Plancitoxins, lethal factors from the crown-of-thorns starfish *Acanthaster planci*, are deoxyribonucleases II. Toxicon **44**:499-506.

111. **Weis MV.** 2008. Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. J. Exp. Biol. **211**:3059-3066.

112. **Ryu J, Ha E, Lee W**. 2010. Innate immunity and gut–microbe mutualism in *Drosophila*. Dev. Comp. Immunol. **34**:369-376.

113. **Boettcher K, Ruby E, McFall-Ngai MJ.** 1996. Bioluminescence in the symbiotic squid *Euprymna scolopes* is controlled by a daily biological rhythm. J. Comp. Phisiol. A. **175**: 65-73.

114. **Shirasu-Hiza MM, Dionne MS, Pham LN, Ayres JS, Schneider DS**. 2007. Interactions between circadian rhythm and immunity in *Drosophila melanogaster*. Curr. Biol. **17**:R353-R355.

115. **Binuramesh C, Michael RD.** 2011. Diel variations in the selected serum immune parameters in Oreochromis mossambicus. Fish Shellfish Immunol. **30**:824-829.

116. **Pham LN, Dionne MS, Shirasu-Hiza M, Schneider DS.** 2007. A specific primed immune response in *Drosophila* is dependent on phagocytes. PLoS Pathogens **3**:e26.

117. **Benkendorff K, Davis AR, Bremner JB.** 2001. Chemical defense in the egg masses of benthic invertebrates: an assessment of antibacterial activity in 39 mollusks and 4 polychaetes. J. Invertebr. Pathol. **78**:109-18.

118. **Fields W.** 1962. The structure, development, food relations, reproduction and life history of the squid *Loligo opalescens* Berry. Fish Bulletin 131. State of California Department of Fish and Game, Sacramento, CA. 1962.

119. **Arnold JM, Singley CT, Williams-Arnold LD.** 1972. Embryonic development and post-hatching survival of the sepiolid squid *Euprymna scolopes* under laboratory conditions. The Veliger **14**:361-364.

120. **Montgomery M, McFall-Ngai MJ.** 1993. Embryonic development of the light organ of the sepiolid squid *Euprymna scolopes* Berry. Biol. Bull.

121. **Bloodgood RA.** 1977. The squid accessory nidamental gland: ultrastructure and association with bacteria. Tissue & Cell 9:197-208.

122. **Lum-Kong A, Hastings T.** 1992. The accessory nidamental glands of *Loligo forbesi* (Cephalopoda: Loliginidae): characterization of symbiotic bacteria and preliminary experiments to investigate factors contribute to sexual maturation. J. Zool. **228**: 395-403.

123. **Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG.** 2009. A single regulatory gene is sufficient to alter bacterial host range. Nature **458**:215-8.

124. **Koch EJ, Miyashiro T, McFall-Ngai MJ, Ruby EG.** 2013. Features governing symbiont persistence in the squid-vibrio association. Mol. Ecol. **23:** 1624-1634

125. **Reasoner DJ, Geldreich EE.** 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. **49:**1-7.

126. **Reynolds ES**. 1963. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. J. Cell Biol. **17:**208-12.

127. **Huber T, Faulkner G, Hugenholtz P.** 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics **20**:2317-2319.

128. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. **72**:5069-5072.

129. **Gomez-Alvarez V, Teal TK, Schmidt TM.** 2009. Systematic artifacts in metagenomes from complex microbial communities. ISME J. **3**:1314-1317.

130. **Meyer F, Paarmann D, D'Souza M, Olson R, Glass E, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards R.** 2008. The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics **9**:386.

131. **Loy A, Loy E, Arnold R, Rol, Tischler P, Arnold, Rattei T, Wagner M, Horn M.** 2008. probeCheck--a central resource for evaluating oligonucleotide probe coverage and specificity. Environ. Microbiol. **10**:2894-8.

132. **Jensen T, Sicko L.** 1971. Fine structure of poly-β-hydroxybutyric acid granules in a blue-green alga, *Chlorogloea fritschii*. J. of Bacteriol. **106**: 683-686

133. **Mandon K, on, Michel-Reydellet N, Encarnación S, Kaminski PA, Leija A, Cevallos MA, Elmerich C, Mora J.** 1998. Poly-beta-hydroxybutyrate turnover in *Azorhizobium caulinodans* is required for growth and affects nifA expression. J. Bacteriol. **180**:5070-6.

134. **Petroni G, Spring S, Schleifer K.** 2000. Defensive extrusive ectosymbionts of *Euplotidium* (Ciliophora) that contain microtubule-like structures are bacteria related to *Verrucomicrobia*. Proc. Nat. Acad. Sci. U.S.A. **97**:1813-1817

135. **Romero-Pérez GA, Ominski KH, McAllister TA, Krause DO.** 2011. Effect of environmental factors and influence of rumen and hindgut biogeography on bacterial communities in steers. Appl. Environ. Microbiol. **77**:258-68.

136. **Vandekerckhove TT, Willems A, Gillis M, Coomans A.** 2000. Occurrence of novel verrucomicrobial species, endosymbiotic and associated with parthenogenesis in

Xiphinema americanum-group species (Nematoda, Longidoridae). Int. J. Syst. Evol. Microbiol. **6**:2197-205.

137. **Yildirim S, Yeoman CJ, Sipos M, Torralba M, Wilson BA, Goldberg TL, Stumpf RM, Leigh SR, White BA, Nelson KE.** 2010. Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. PLoS ONE **5**:e13963.

138. **Kaufman M, Ikeda Y, Patton C, Dykhuizen G.** 1998. Bacterial symbionts colonize the accessory nidamental gland of the squid *Loligo opalescens* via horizontal transmission. Biol. Bull. **194**: 36-43

139. **Hanlon RT, Claes MF, Ashcraft SE, Dunlap PV.** 1997. Laboratory culture of the sepiolid squid *Euprymna scolopes*: a model system for bacteria-animal symbiosis. Biol. Bull. **192**:364-374.

140. **Branden C, Richard A, Lemaire J.** 1979. La glande nidamentaire accessoire de Sepia officinalis L.: analyses biochimiques des pigments des bactéries symbiotiques. Annales Soc. r. Zool. Belg. **108**: 123-139.

141. **Yoon J, Yasumoto-Hirose M, Katsuta A, Sekiguchi H, Matsuda S, Kasai H, Yokota A.** 2007. Coraliomargarita akajimensis gen. nov., sp. nov., a novel member of the phylum "Verrucomicrobia" isolated from seawater in Japan. 5. Int. J. Syst. Evol. Microbiol. **57**:959-963.

142. **Derrien M, Vaughan EE, Plugge CM, de Vos WM.** 2004. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. Int. J. Syst. Evol. Microbiol. 54:1469-76.

143. **Sakai, Sakai T, T., et, Kawai T, Kato I.** 2004. Isolation and characterization of a fucoidan-degrading marine bacterial strain and its fucoidanase. Mar. Biotechnol. **6**:335-346.

144. **Collins AJ, Nyholm SV.** 2011. Draft genome of *Phaeobacter gallaeciensis* ANG1, a dominant member of the accessory nidamental gland of Euprymna scolopes. J. Bacteriol. **193**:3397-8.

145. **Ratcliff WC, Kadam SV, Denison RF.** 2008. Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. *FEMS Microbiol. Ecol.* **65**:391-399.

146. **Barbieri E, Barry K, Child A, Wainwright N.** 1997. Antimicrobial activity in the microbial community of the accessory nidamental gland and egg cases of *Loligo pealei* (Cephalopoda: Loliginidae). Biol. Bull. **193**:275-276.

147. **Gomathi P, Nair Jr, Sherief P.** 2010. Antibacterial activity in the accessory nidamental gland extracts of the Indian squid, *Loligo duvauceli* Orbigny. Indian. J. Mar. Sci. **39**:100-104.

148. **Cude WN, Mooney J, Tavanaei AA, Hadden MK, Frank AM, Gulvik CA, May AL, Buchan A.** 2012. Production of the antimicrobial secondary metabolite indigoidine contributes to competitive surface colonization by the marine roseobacter *Phaeobacter sp.* strain Y4I. Appl. Environ. Microbiol. **78**:4771-80.

149. **Lane DJ.** 1991. 16S/23S rRNA Sequencing, pp. 115-175. In Stakebrant, E, Goodfellow, M (eds.), Nucleic acid techniques in bacterial systematics. John Wiley and Sons, New York, NY.

150. **Manz W, Amann R, Ludwig W, Vancanneyt M, Schleifer KH.** 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum *Cytophaga-Flavobacter-Bacteroides* in the natural environment. Microbiol. **142**:1097-106.

151. **Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA.** 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl. Environ. Microbiol. **56**:1919-25.

152. **Wallner G, Amann R, Beisker W.** 1993. Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry 14:136-43.

153. **Daims H, Brühl A, Amann R, Schleifer K, Wagner M.** 1999. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. Syst. Appl. Microbiol. 22:434-444.

154. **Neef A.** Anwendung der in situ Einzelzell - Identifizierung von Bakerien zur Populationsanalyse in komplexen mikrobiellen blozonosen. PhD Thesis. Technical University, Munich, Germany.

155. **Giuliano L, de M D, de E D, Höfle M, Yakimov M.** 1999. Identification of culturable oligotrophic bacteria within naturally occurring bacterioplankton communities of the Ligurian Sea by 16S rRNA sequencing and probing. Microb. Ecol. **37**:77-85.

156. **Case RJ, Longford SR, Alex, Campbell AH, Campbell RH, Low A, Tujula N, Steinberg PD, Kjelleberg S.** 2011. Temperature induced bacterial virulence and bleaching disease in a chemically defended marine macroalga. Environ. Microbiol. **13**:529-37.

157. **Gil-Turnes MS, Hay ME, Fenical W.** 1989. Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. Science **246**:116-8.

158. **Brinkhoff T, Bach G, Heidorn T, Liang L, Schlingloff A, Simon M.** 2004. Antibiotic production by a *Roseobacter* clade-affiliated species from the german Wadden Sea and its antagonistic effects on indigenous isolates. Appl. Environ. Microbiol. **70**:2560-2565.

159. **Seyedsayamdost MR, Case RJ, Kolter R, Clardy J.** 2011. The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis.* Nat Chem **3**:331-5.

160. **Dogs M, Göker M, Teshima H, Teshima H, Petersen J, Petersen J, Fiebig A, Fiebig A, Chertkov O, Dalingault H, Chen A, Pati A, Goodwin LA, Chain P, Detter JC, Ivanova N, Lapidus A, Rohde M, Gronow S, Kyrpides NC, Woyke T, Simon M, Göker M, Klenk H, Brinkhoff T.** 2013. Genome sequence of *Phaeobacter daeponensis* type strain (DSM 23529(T)), a facultatively anaerobic bacterium isolated from marine sediment, and emendation of *Phaeobacter daeponensis*. Standards in Genomic Sciences **9**:142-59.

161. **Epel D.** 2004. Environmental effects on anti-microbial activity of bacterial symbionts in the reproductive system of squid. UCSD Research Final Reports. **89**:1-6.

162. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R.** 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. **42**:D206-14.

163. **Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T.** 2013. antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. **41**:W204-12.

164. **van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP.** 2013. BAGEL3: Automated identification of genes encoding bacteriocins and (non)bactericidal posttranslationally modified peptides. Nucleic Acids Res. **41**:W448-53.

165. **Rice P, Longden I, Bleasby A.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends. Genet. **16**:276-7.

166. **Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE, Howard EC, King E, Oakley CA, Reisch CR, Rinta-Kanto JM, Sharma S, Sun S, Varaljay V, Vila-Costa M, Westrich JR, Moran MA.** 2010. Genome characteristics of a generalist marine bacterial lineage. ISME J **4**:784-798.

167. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792-1797.

168. **Darriba D, Taboada GL, Doallo R, Posada D**. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods **9**:772-772.

169. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. **59**:307-321.

170. **Richter M, Rossello-Mora R.** 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc. Nat. Acad. Sci. U.S.A **106**:19126-19131.

171. **Whistler CA, Ruby EG.** 2003. GacA regulates symbiotic colonization traits of *Vibrio fischeri* and facilitates a beneficial association with an animal host. J. Bacteriol. **185**:7202-7212.

172. **McMillan DGG, Velasquez I, Nunn BL, Goodlett DR, Hunter KA, Lamont I, Sander SG, Cook GM.** 2010. Acquisition of iron by alkaliphilic bacillus species. Appl. Environ. Microbiol. **76**:6955-6961.

173. **Schwyn B, Neilands JB**. 1987. Universal chemical assay for the detection and determination of siderophores. Anal. Biochem. **160**:47-56.

174. **Arnow LE**. 1937. Colorimetric determination of the components of 3,4-dihydroxyphenylalaninetyrosine mixtures. J. Biol. Chem. **118:** 531-537.

175. **Csaky TZ**. 1948. On the estimation of bound hydroxylamine in biological materials. Act. Chem. Scand. **2**: 450-454.

176**. Cha C, Gao P, Chen Y, Shaw P.** 1998. Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. Mol. Plant Microbe Interact. **11**: 1119-1129.

177**. Ravn L, Christensen AB, Molin S, Givskov M, Gram L.** 2001. Methods for detecting acylated homoserine lactones produced by Gram-negative bacteria and their application in studies of AHL-production kinetics. J. Microbial.l Met. **44**:239-51.

178. **Chilton MD, Currier TC, Farr SK, Farrand SK, Bendich AJ, Gordon MP, Nester EW.** 1974. *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. Proc. Natl. Acad. Sci. U.S.A. **71**:3672-6.

179. **Moran MA, Buchan A, González JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N.** 2004. Genome sequence of Silicibacter pomeroyi reveals adaptations to the marine environment. Nature **432**:910-913.

180. **Berger A, Dohnt K, Tielen P, Jahn D, Becker J, Wittmann C.** 2014. Robustness and Plasticity of Metabolic Pathway Flux among Uropathogenic Isolates of Pseudomonas aeruginosa. PLoS ONE **9**:e88368-14.

181. **Fuchs G. 2009. Oxidation of Organic Compounds,** pp. 187-223. In Lengeler, JW, Drews, G, Schlegel, HG (eds.). Biology of the Prokaryotes. Georg Thieme Verlag, Sttutgart, Germany.

182. **Beyersmann P, Chertkov O.** 2013. Genome sequence of *Phaeobacter caeruleus* type strain (DSM 24564 T), a surface-associated member of the marine *Roseobacter* clade. Stand. Genomic. Sci. **8**: 403-419

183. **Riedel T, Teshima H, Petersen J, Fiebig A, Davenport K, Daligault H, Erkkila T, Gu W, Munk C, Xu Y, Chen A, Pati A, Ivanova N, Goodwin LA, Chain P, Detter JC, Rohde M, Gronow S, Kyrpides NC, Woyke T, Göker M, Brinkhoff T, Klenk H.** 2013. Genome sequence of the *Leisingera aquimarina* type strain (DSM 24565(T)), a member of the marine *Roseobacter* clade rich in extrachromosomal elements. Stand. Genomic Sci. **8**:389-402.

184. **Buddruhs N, Chertkov O, Petersen J, Fiebig A, Chen A, Pati A, Ivanova N, Lapidus A, Goodwin LA, Chain P, Detter JC, Gronow S, Kyrpides NC, Woyke T, Göker M, Brinkhoff T, Klenk H.** 2013. Complete genome sequence of the marine methyl-halide oxidizing *Leisingera methylohalidivorans* type strain (DSM 14336(T)), a representative of the *Roseobacter* clade. Stand. Genomic Sci. **9**:128-41.

185. **Murdoch SL, Trunk K, English G, Fritsch MJ, Pourkarimi E, Coulthurst SJ.** 2011. The opportunistic pathogen *Serratia marcescens* utilizes type VI secretion to target bacterial competitors. J. Bacteriol. **193**:6057-6069.

186. **Hood RD, Singh P, Hsu F, GUvener T, Carl MA, Trinidad RRS, Silverman JM, Ohlson BB, Hicks KG, Plemel RL, Li M, S, Schwarz R, Wang WY, Merz AJ, Goodlett DR, Mougous JD, Güvener T, Schwarz S.** 2010. A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. Cell Host & Microbe **7**:25-37.

187. **Wu H, Chung P, Shih H, Wen S, Lai E**. 2008. Secretome analysis uncovers an Hcp-family protein secreted via a type VI secretion system in *Agrobacterium tumefaciens*. J. Bacteriol. **190**:2841-2850.

188. **Bladergroen MR, Badelt K, Spaink HP.** 2003. Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion. Mol. Plant Microbe Interact. **16**:53-64.

189. **Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ.** 2006. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc. Natl. Acad. Sci. **103**:1528-33.

190. **Williams SG, Varcoe LT, Attridge SR, Manning PA.** 1996. *Vibrio cholerae* Hcp, a secreted protein coregulated with HlyA. Infect. Immun. **64**:283-9.

191. **Gibbs KA, Urbanowski ML, Greenberg EP.** 2008. Genetic determinants of self identity and social recognition in bacteria. Science **321**:256-259.

192. **Geng H, Bruhn JB, Nielsen KF, Gram L, Belas R.** 2008. Genetic dissection of tropodithietic acid biosynthesis by marine roseobacters. Appl. Environ. Microbiol. **74**:1535-45.

193. **Martens T, Gram L, Grossart H, Kessler D, Müller R, Simon M, Wenzel SC, Brinkhoff T.** 2007. Bacteria of the *Roseobacter* clade show potential for secondary metabolite production. Microb. Ecol. **54**:31-42.

194. **Antunes LCM, Schaefer AL, Ferreira RBR, Qin N, Stevens AM, Ruby EG, Greenberg EP**. 2007. Transcriptome analysis of the *Vibrio fischeri* LuxR-LuxI regulon. J. Bacteriol. **189**:8387-8391.

195. **Rosemeyer V, Michiels J, Verreth C, J V, Vanderleyden J.** 1998. luxI- and luxR-homologous genes of *Rhizobium etli* CNPAF512 contribute to synthesis of autoinducer molecules and nodulation of *Phaseolus vulgaris*. J. Bacteriol. 180:815-21.

196. **Daniels R, de Vos DE, Desair J, Raedschelders G, Luyten E, Rosemeyer V, Verreth C, Schoeters E, Jos V, Vanderleyden J, erleyden, Michiels J.** 2002. The cin quorum sensing locus of *Rhizobium etli* CNPAF512 affects growth and symbiotic nitrogen fixation. J. Biol. Chem. **277**:462-8.

197. **Andrews SC, Robinson AK, Rodríguez-Quiñones F.** 2003. Bacterial iron homeostasis. FEMS Microbiol. Rev. **27**:215-37.

198. **Brickman TJ, Armstrong SK.** 1996. The ornithine decarboxylase gene odc is required for alcaligin siderophore biosynthesis in *Bordetella spp*.: putrescine is a precursor of alcaligin. J. Bacteriol. **178**: 54-60.

199. **Crosa JG, Walsh CT.** 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. Microbiol. Mol. Biol. Rev. **66**:223-249.

200. **Ciche TA, Blackburn M, Carney JR, Ensign JC**. 2003. Photobactin: a catechol siderophore produced by *Photorhabdus luminescens*, an entomopathogen mutually associated with *Heterorhabditis bacteriophora* NC1 nematodes. Appl. Environ. Microbiol. **69**: 4706-4713.

201. **Raymond KN, Dertz EA, Kim SS**. 2003. Enterobactin: an archetype for microbial iron transport. Proc. Nat. Acad. Sci. U.S.A. **100**:3584-3588.

202. **Goetz DH, Holmes MA, Borregaard N, Bluhm ME, Raymond KN, Rol, Strong RK, Strong K.** 2002. The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. Mol. Cell **10**:1033-43.

203. **Abergel RJ, Clifton MC, Pizarro JC, Warner JA, Shuh DK, Rol, Strong RK, Strong K, Raymond KN.** 2008. The siderocalin/enterobactin interaction: a link between mammalian immunity and bacterial iron transport. J. Am. Chem. Soc. **130**:11524-34.

204. **Ong DST, Ong DST, Wang L, Zhu Y, Ho B, Ding JL, Ho B.** 2005. The response of ferritin to LPS and acute phase of Pseudomonas infection. *J. Endotoxin Res.* **11**:267-280.

205. **Wang D, Kim BY, Lee KS, Yoon HJ, Cui Z, Lu W, Jia JM, Kim DH, Sohn HD, Jin BR.** 2009. Molecular characterization of iron binding proteins, transferrin and ferritin heavy chain subunit, from the bumblebee *Bombus ignitus*. Comp. Biochem. Physiol. B, Biochem. Mol. Biol. **152**:20-27.

206. **Buchon N, Broderick NA, Poidevin M, Praderv S, Lemaitre B, Pradervand S.** 2009. Drosophila Intestinal Response to Bacterial Infection: Activation of Host Defense and Stem Cell Proliferation *Cell Host & Microbe* **5**:200-211.

207. **Garver LS, Wu J, Wu LP.** 2006. The peptidoglycan recognition protein PGRP-SC1a is essential for Toll signaling and phagocytosis of *Staphylococcus aureus* in *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. **103**:660-5.

208. **Poltorak A, He X, Smirnova I, Liu MY, van Huffel C, X du, Birdwell D, Alejos E, Silva M, Galanos C, Freudenberg M, Ricciardi-Castagnoli P, Layton B, Beutler B.** 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science **282**:2085-8.

209. **Takeda K.** 2004. Toll-like receptors in innate immunity. Int. Immunol. **17**:1-14.

210. **Yu Y, Yuan S, Yu Y, Huang H, Feng K, Pan M, Huang S, Dong M, Chen S, Xu A.** 2007. Molecular and biochemical characterization of galectin from amphioxus: primitive galectin of chordates participated in the infection processes. Glycobiology 17:774-83.

211. **Dziarski R.** 2003. Defect in neutrophil killing and increased susceptibility to infection with nonpathogenic gram-positive bacteria in peptidoglycan recognition protein-S (PGRP-S)-deficient mice. Blood **102**:689-697.

212. **Won B.** 2009. Characterizing the Role of Phaeobacter in the Mortality of the Squid, *Euprymna scolopes*. Honors Thesis. University of Connecticut, Storrs, CT.

213. **Vaishnava S, Behrendt CL, Ismail AS, Eckmann L, Hooper LV.** 2008. Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. Proc. Nat. Acad. Sci. **105**:20858-20863.

214. **Fraune S, Augustin R, Anton-Erxleben F, Wittlieb J, Gelhaus C, Klimovich BV, Samoilovich MP, Bosch TCG.** 2010. In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. Proc. Nat. Acad. Sci. U.S.A. **107**:18067-18072.

215. **Fukui Y, Abe M, Kobayashi M, Ishihara K, Oikawa H, Yano Y, Satomi M.** 2012. *Maritalea porphyrae* sp. nov., isolated from a red alga (*Porphyra yezoensis*), and transfer of *Zhangella mobilis* to *Maritalea mobilis* comb. nov. Int. J. Syst.Evol. Microbiol. **62**:43-8.

216. **Mao Y, Wang J, Zhang Z, Ding S, Su Y.** 2010. Cloning, mRNA expression, and recombinant expression of peptidoglycan recognition protein II gene from large yellow croaker (*Pseudosciaena crocea*). *Mol. Biol. Rep.* **37**:3897-3908.

217. **Studier FW.** 2005. Protein production by auto-induction in high-density shaking cultures. Prot. Exp. and Purif. **41**:207-234.