Colonization State Influences the Hemocyte Proteome in a Beneficial Squid–Vibrio Symbiosis*[S]

Tyler R. Schleicher‡, Nathan C. VerBerkmoes§¶, Manesh Shah‖, and Spencer V. Nyholm‡**

The squid *Euprymna scolopes* and the luminescent bacterium *Vibrio fischeri* form a highly specific beneficial light organ symbiosis. Not only does the host have to select *V. fischeri* from the environment, but it must also prevent subsequent colonization by non-symbiotic microorganisms. Host macrophage-like hemocytes are believed to play a role in mediating the symbiosis with *V. fischeri*. Previous studies have shown that the colonization state of the light organ influences the host's hemocyte response to the symbiont. To further understand the molecular mechanisms behind this process, we used two quantitative mass-spectrometry-based proteomic techniques, isobaric tags for relative and absolute quantification (iTRAQ) and label-free spectral counting, to compare and quantify the adult hemocyte proteomes from colonized (sym) and uncolonized (antibiotic-treated/cured) squid. Overall, iTRAQ allowed for the quantification of 1,024 proteins with two or more peptides. Thirty-seven unique proteins were determined to be significantly different between sym and cured hemocytes (p value < 0.05), with 20 more abundant proteins and 17 less abundant in sym hemocytes. The label-free approach resulted in 1,241 proteins that were identified in all replicates. Of 185 unique proteins present at significantly different amounts in sym hemocytes (as determined by spectral counting), 92 were more abundant and 93 were less abundant. Comparisons between iTRAQ and spectral counting revealed that 30 of the 37 proteins quantified via iTRAQ exhibited trends similar to those identified by the label-free method. Both proteomic techniques mutually identified 16 proteins that were significantly different between the two groups of hemocytes (p value < 0.05). The presence of *V. fischeri* in the host light organ influenced the abundance of proteins associated with the cytoskeleton, adhesion, lysosomes, proteolysis, and the innate immune response. These data provide evidence that colonization by *V. fischeri* alters the hemocyte proteome and reveals proteins that may be important for maintaining host–symbiont specificity. ***Molecular & Cellular Proteomics 13: 10.1074/mcp.M113.037259, 2673–2686, 2014.***

Animals must have mechanisms that allow them to differentiate between their normal microbiota and pathogenic microorganisms. In order to recognize microorganisms, vertebrates can utilize adaptive immunity, which is mediated by the ability to produce antibodies to specific antigens. However, invertebrates lack this antibody-based component, and immunological memory is poorly understood in these animals (1). Instead, invertebrates rely on components of the innate immune system to interact with microorganisms. Long recognized for its ability to remove pathogens, the innate immune system has more recently been studied for interactions that foster benign and beneficial symbioses (2). Factors such as reactive oxygen and nitrogen species, antimicrobial peptides, and complement-like proteins have a function in maintaining populations of beneficial bacteria in a number of animal hosts (2). In addition, invertebrates have a cellular component to their innate immune system, specifically, macrophage-like cells that can bind and phagocytose bacteria (3–5). These cells utilize pattern recognition receptors, which can detect microbe-associated molecular patterns such as peptidoglycan and lipopolysaccharide, to activate an immune response (2). A growing body of evidence from several invertebrates suggests that these phagocytic cells are capable of responses to specific bacteria (3, 6, 7).

The Hawaiian bobtail squid, *Euprymna scolopes*, is used as a model organism to study the role of the innate immune response in a beneficial symbiosis (2, 8). Upon hatching, the host establishes a highly specific binary association with the bacterium *Vibrio fischeri*, which colonizes a specialized light organ (9). Inside this organ *V. fischeri* is extracellular and maintained within epithelium-lined crypt spaces. Although the light organ is open to the environment, non-symbiotic bacteria are excluded, and *V. fischeri* remains the sole symbiont for the

From the ‡Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut, 06269; §Chemical Biology Division, New England Biolabs Inc., Ipswich, Massachusetts, 01938; ‖Graduate School of Genome Science and Technology, University of Tennessee, Knoxville, Tennessee, 37996

Received December 20, 2013, and in revised form, July 16, 2014
Published, MCP Papers in Press, July 18, 2014, DOI 10.1074/mcp.M113.037259
Author contributions: T.R.S. and S.V.N. designed research; T.R.S., N.C.V., and S.V.N. performed research; N.C.V. contributed new reagents or analytic tools; T.R.S., N.C.V., M.S., and S.V.N. analyzed data; T.R.S., N.C.V., and S.V.N. wrote the paper.

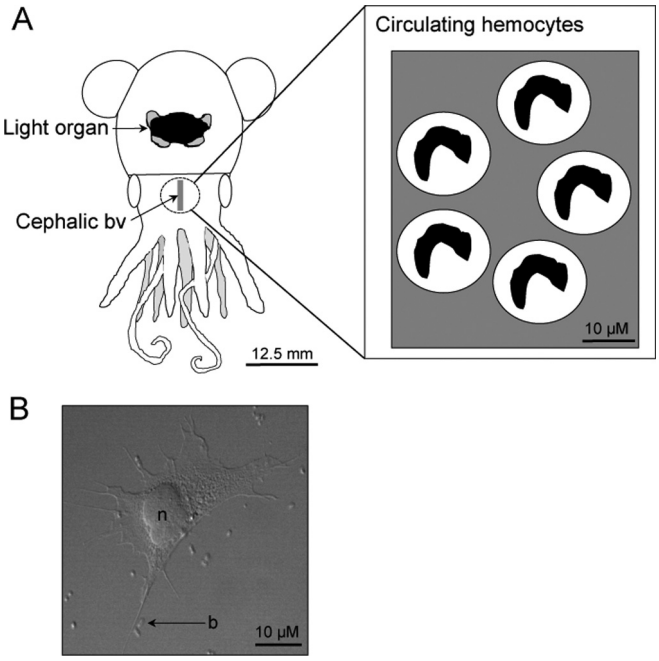

FIG. 1. ***Euprymna scolopes* has one type of blood cell, the macrophage-like hemocyte.** *A*, the dorsal view of *E. scolopes* reveals the location of the light organ and the cephalic blood vessel (bv). The hemocytes circulate through the vasculature of the squid (scale: 12.5 mm). *B*, a differential interference contrast image of an adult *E. scolopes* hemocyte interacting with *Vibrio fischeri* cells (b, bacteria; n, nucleus; scale, 10 μm).

duration of the animal's life. Although the mechanisms behind this specificity are not completely understood, studies suggest that macrophage-like hemocytes play a role (2, 3, 8, 10).

Hemocytes are the only type of blood cell found in *E. scolopes* (Fig. 1). Upon initiation of the symbiosis, these hemocytes traffic to the juvenile light organ in response to *V. fischeri* (11). Within the colonized juvenile light organ crypt spaces, hemocytes with internalized bacteria have been observed (12). It is unclear whether these bacteria are *V. fischeri* or non-symbiotic interlopers that have entered the organ from the environment. However, the hemocytes of an adult squid have never been visualized with engulfed bacteria in the light organ after the symbiosis has had time to mature, even among the dense population of *V. fischeri* in the crypt spaces. A previous binding study revealed that adult hemocytes adhere to the symbiont significantly less than related non-symbiotic bacteria (3). Removal of the symbiont from the light organ with antibiotics leads to hemocytes that adhere to *V. fischeri* in significantly greater numbers, suggesting that colonization induces host immune tolerance of the symbiont (3).

Although the molecular mechanisms behind this change in hemocyte response to *V. fischeri* remain poorly characterized, a transcriptome and proteome identified hemocyte genes and proteins that may be important to the symbiosis (10). Quantitative PCR revealed that nitric oxide synthase, complement component C3, and peptidoglycan recognition protein 5 (PGRP 5)[1] were differentially expressed between the hemocytes from colonized (sym) and cured squid (10). In order to further understand the effect of light organ colonization on hemocyte function, we used two quantitative MS-based proteomic techniques, isobaric tags for relative and absolute quantification (iTRAQ) and label-free spectral counting, to compare the proteomes of sym and cured hemocytes. Proteins that showed significant differences in abundance in sym hemocytes included those involved with the immune response, adhesion, the cytoskeleton, and lysosomal processes. These data, along with the results from previous studies, suggest that colonization by *V. fischeri* alters the host's immune response by influencing the hemocyte proteome to favor tolerance of the symbiont.

EXPERIMENTAL PROCEDURES

Animal Collection and Maintenance—Adult *E. scolopes* were caught within the shallow sand flats off of Oahu, HI, by dip net. Hemocyte samples not collected immediately in Hawaii were obtained from animals maintained at the University of Connecticut in artificial seawater (Instant Ocean) at 23 °C on an approximate 12-h light/12-h dark cycle (10, 13). All animals were allowed to acclimate for at least 48 h under laboratory conditions prior to sample collection.

Curing Experiments (Symbiont Removal)—*V. fischeri* was cured from the light organ using a mixture of chloramphenicol (20 μg/ml; Sigma Aldrich, St. Louis, MO) and gentamicin (20 μg/ml; Affymetrix/USB, Cleveland, OH) for 5 consecutive days (3, 10). We confirmed curing by homogenizing the symbiont-containing central core of the light organ and plating the homogenate on seawater tryptone agar (3, 10). Plates with central core homogenates from cured hosts contained no detectable *V. fischeri* colonies after an overnight incubation at 28 °C. Hemocytes from cured hosts were collected and prepared as described below.

Hemocyte Collection and Protein Extraction—Squid hemocytes were collected from adult *E. scolopes* as previously described (3, 10, 14). Approximately 50 to 100 μl of hemolymph (~5,000 hemocytes per microliter (3)) was extracted from the cephalic blood vessel using a sterile 1-ml syringe with a 28-gauge needle. Freshly collected hemocytes were washed in 100 to 200 μl of Squid Ringer's solution (530 mM NaCl, 10 mM KCl, 25 mM $MgCl_2$, 10 mM $CaCl_2$, and 10 mM HEPES buffer, pH 7.5) two times to remove excess hemolymph. Hemocytes collected from untreated/colonized squid were designated as "sym," and hemocytes from antibiotic-treated/cured squid were referred to as "cured."

Sym hemocyte samples ($n = 8$) and cured hemocyte samples ($n = 8$) were resuspended in 400 μl of radioimmune precipitation assay buffer (50 mM Tris, 150 mM NaCl, 1% Triton, 0.5% sodium deoxycholate, 0.1% SDS, pH 7.5), pooled, and homogenized with a pestle, resulting in two samples, sym and cured. The lysates were centrifuged (Eppendorf 5810 R, 14,000 rpm, 4 °C, 30 min), and the protein concentration of the supernatants was quantified spectrophotometrically using the *RC DC* protein assay (Bio-Rad, Hercules, CA). Approximately 300 μg of protein were obtained from sym hemocytes, and 500 μg from cured hemocytes.

[1] The abbreviations used are: PGRP, peptidoglycan recognition protein; Sym, symbiotic; iTRAQ, isobaric tags for relative and absolute quantification; mPBS, marine phosphate-buffered saline; FASP, filter-assisted sample preparation; NSAF, normalized spectral abundance factor.

Immunocytochemistry—Hemocytes were prepared for immunocytochemistry as previously described, with minor modifications (15). Approximately 50 μl of hemolymph per squid was adjusted to a total volume of 300 μl with Squid Ringer's and then evenly distributed among three wells (containing glass coverslips submerged in 1 ml of Squid Ringer's in a 12-well tissue culture plate (~8.3×10^4 hemocytes/slide)). Hemocytes were allowed to adhere to glass coverslips for 30 min (all steps were carried out at room temperature). After attachment, the hemocytes were rinsed two times in Squid Ringer's to remove excess tissue or cellular debris and then fixed with 4% paraformaldehyde in Squid Ringer's for 30 min. After fixation, the hemocytes were washed four times with marine phosphate-buffered saline (mPBS), pH 7.4 (50 mM sodium phosphate, 0.45 M NaCl), for 10 min each time. Fixed hemocytes were permeabolized with 1% Triton X-100, mPBS, pH 7.4, for 30 min. The cells were blocked (mPBS, pH 7.4, 1% Triton X-100, 1% goat serum, 0.5% BSA) for 1 h. Hemocytes were then exposed to anti-cathepsin L2 (donated by M. McFall-Ngai, University of Wisconsin-Madison) at a 1:1,000 dilution in block overnight. The primary antibody was then removed, and the hemocytes were washed four times in 1% Triton X-100, mPBS, pH 7.4, for 5 min each time and then blocked for an additional hour. The hemocytes were then incubated with an Alexa Fluor 488 goat anti-rabbit secondary antibody (Invitrogen, Grand Island, NY) at a 1:1,000 dilution in fresh block for 2 h in the dark. Hemocytes were washed four times in 1% Triton X-100, mPBS, pH 7.4, for 5 min each time and then counterstained using a nuclear stain, DRAQ5 (2.5 μM; Thermo Scientific, Rockford, IL), in 1% Triton X-100, mPBS, pH 7.4, overnight in the dark. Finally, the hemocytes were mounted onto glass slides with Vectashield (Vector Laboratories, Burlingame, CA), sealed with nail polish, and imaged using an A1R confocal microscope equipped with NIS-Elements software (v. 4.13). Z-sections (0.225 μm each) were collected for more than 10 hemocytes per squid combined from more than two microscopic fields (three sym squid and two cured squid). Maximum-intensity projections and fluorescence measurements were made using FIJI as previously described (16, 17).

iTRAQ Methodology

Protein Digestion and iTRAQ Labeling—The pooled protein stocks collected from sym and cured hemocytes were split into two samples as experimental replicates (Fig. 2). Proteins were precipitated using a methanol and chloroform protocol as previously described (18). The resulting pellet was resuspended in 0.5 M triethylammonium bicarbonate, 2% SDS. The proteins were reduced using 50 mM tris-(2-carboxyethyl) phosphine (Thermo Scientific, Rockford, IL) at 60 °C for 1 h and alkylated with 200 mM methyl methanethiosulfonate (Thermo Scientific) at room temperature for 10 min. Samples were digested with trypsin (Promega, Madison, WI) at a 1:10 ratio (enzyme:protein) at 37 °C overnight (18 to 20 h). iTRAQ labeling was performed on 100 μg of protein from each replicate per condition using the reagents from a 4-plex iTRAQ kit (AB SCIEX, Framingham, MA). The labeled samples were then pooled.

Strong Cation Exchange—The labeled hemocyte protein samples were acidified with 1 M phosphoric acid to a pH less than 3.0 and then separated on a Hewlett Packard 1090 HPLC system fitted with a polySulfoethylA column (The Nest Group, Southborough, MA). At 0.5 ml/min of buffer A (10 mM KH_2PO_4, pH 3.0, 25% acetonitrile), a gradient of 0% to 100% buffer B (10 mM KH_2PO_4, 1 M NaCl, pH 3.0, 25% acetonitrile) was established over 120 min. Fractions were collected at 1-min intervals. The broad, unresolved A214 peak was pooled into 10 fractions according to absorbance. An additional cleanup step was performed for each fraction with a C18 MacroSpin column according to the manufacturer's protocol (The Nest Group).

LC-MS/MS—iTRAQ samples were analyzed via LC-MS/MS at Yale University's W.M. Keck Biotechnology Resource Laboratory. Three micrograms of each strong cation exchange fraction ($n = 10$) were separated and analyzed on a Waters nanoAcquity UPLC system equipped with a 5600TripleTOF (AB SCIEX) fitted with a Nanospray III source (AB SCIEX) and a pulled quartz tip as the emitter (New Objectives, Woburn, MA) as previously described, with minor modifications (10). For trapping, a flow rate of 5 μl/min with 99% buffer A (100% water, 0.1% formic acid) was maintained for 1 min using a Waters Symmetry® C18 180 μm × 20 mm trap column. A 1.7-μm, 75 μm × 150 mm nanoAcquity™ UPLC™ column (at 45 °C) was used for peptide separation. At a flow rate of 500 nl/min, a 161-min linear gradient was maintained with buffer A and buffer B (100% CH_3CN, 0.075% formic acid). Initial conditions consisted of 95% buffer A and 5% buffer B. Conditions reached 60% A, 40% B at 160 min and 15% A, 85% B at 161 min. Mass spectrometer settings included an ion spray voltage of 2.2 kV, a curtain gas of 20 PSI, and a sweeping collision energy setting of 35 ± 15 eV for collision-induced dissociation. During information-dependent acquisition mode, survey scans were acquired for 250 ms for a mass range of 400–1250 Da. For ions with a charge state of +2 to +5 that exceeded 125 counts per second, up to 20 product ion scans were collected. Four time bins were summed for each scan at a pulser frequency value of 15.420 kHz through monitoring of the 40-GHz multichannel time to digital converter detector with four-anode/channel detection. The total cycle time was fixed at 1.3 s.

Data Analysis—The combined raw MS/MS files (*.wiff) from Analyst TF 1.5.1 were analyzed with the Paragon™ search algorithm of ProteinPilot (version 4.0 (19)). Data were searched against an *E. scolopes* protein sequence database originating from published transcriptomic sequences (34,684 sequences (10, 20, 21)). The transcriptomic sequences were translated, and the longest open reading frame was annotated using BlastP and the NCBI nr database with an e-value cutoff of $1E^{-3}$ (due to the low representation of cephalopod protein sequences in the nr database). The ProteinPilot software determined the mass tolerance of precursor and fragment ions during the calibration of the data. Searching parameters included the 4-plex iTRAQ reagents, trypsin digestion, methyl methanethiosulfonate cysteine alkylation, bias correction, and background correction, in addition to selecting no special factors, biological modifications, or amino acid substitutions. The false discovery analysis conducted by the ProteinPilot software utilized a reversed-sequence decoy database to determine the false discovery rate (supplemental material). Peptides identified by ProteinPilot were filtered using the "auto" setting to include only unique peptides, no missed cleavages, and at least two iTRAQ ions per peptide. To minimize iTRAQ ratios from low-intensity ions, ProteinPilot also required peptides to maintain a signal-to-noise ratio greater than 9 from the combined intensities of the contributing iTRAQ ions. Additionally, each protein quantified by ProteinPilot required two or more peptides. To compensate for any technical variation between iTRAQ labels, the remaining peptides were subject to a Cyclic Loess normalization using a script designed at Yale University's W.M. Keck MS and Proteomics Resource Laboratory (22). Pairwise comparisons of each iTRAQ reporter during normalization revealed that the experimental variation arising from the iTRAQ reagents and methods was around 1.75-fold, and this was used as the fold-change cutoff (supplemental material). The iTRAQ ion areas for labels 114 and 115 (sym) and for labels 116 and 117 (cured) were averaged for each peptide of each protein. These average iTRAQ values were then used to compare the protein levels between sym and cured hemocytes. Proteins were determined to be significantly more or less abundant if they had a p value less than 0.05 (established by a t test of the iTRAQ ion areas from contributing peptides) and a Log2 fold change greater than 0.80 or less than −0.80 (fold change of 1.75). Only proteins with a ProteinPilot unused score of >2.0 (99% confidence level), a significant p value, and a Log2 fold change

meeting the cutoff were considered for further analysis. The after-normalization ProteinPilot summary and iTRAQ fold-change analysis can be found in the supplemental material.

Label-free Spectral Counting Methodology

Protein Digestion—Proteome samples were digested using a modified filter-assisted sample preparation (FASP) technique (Expedeon, Harston, UK). Briefly, ~100 μg of each proteome sample were diluted in 1% SDS and heated at 60 °C for 5 min. Samples were allowed to cool and then were further diluted with 1% SDS, 8 M urea, 10 mM DTT (made in 50 mM Tris buffer, pH 8.0; all buffers above and below were made in this solution) and rocked at room temperature for 45 min. Samples were then transferred onto Expedeon FASP filters and centrifuged at 14,000 × *g* for 15 min. Fresh 8 M urea was added to the filters and centrifuged at 14,000 × *g* for 15 min. Ten microliters of iodoacetamide solution (provided in the FASP kit) and 90 μl of urea solution (no DTT, no SDS) were added to each filter preparation. Samples were vortexed for 1 min, incubated without mixing for 20 min in the dark, and then centrifuged at 14,000 × *g* for 15 min. One hundred microliters of 8 M urea solution was then added to the filters (no DTT, no SDS), and samples were centrifuged at 14,000 × *g* for 15 min. This step was repeated twice and was followed by the addition of 100 μl of a 50 mM ammonium bicarbonate solution (provided with the FASP kit) and centrifugation at 14,000 × *g* for 15 min. The filter was then transferred to a new collection tube for the proteolytic digestion step. Seventy-five microliters of digestion solution was added to the filter and incubated at 37 °C for 18 h with no rocking. The digestion solution contained 75 μl of ammonium bicarbonate solution with the addition of 10 μg of trypsin (Trypsin Ultra, New England Biolabs, Ipswich, MA). The next day peptides were eluted into the clean collection tube via three steps of centrifugation at 14,000 × *g* for 15 min as follows: Step 1, 40 μl of 50 mM ammonium bicarbonate solution; Step 2, 50 μl of 0.5 M sodium chloride solution (provided with the FASP kit); and Step 3, 170 μl of H_2O with formic acid. The final solution was split into three aliquots per sample and frozen at −80 °C until analysis via two-dimensional LC-MS/MS. The eluted solution was ready to load onto a two-dimensional nano-LC column without further purification or desalting.

Two-dimensional LC-MS/MS Analyses—Tryptic digests of sym and cured unlabeled protein were analyzed in technical duplicates via a two-dimensional nano-LC-MS/MS system with a split-phase nano column (RP-SCX-RP (23)) on a QExactive mass spectrometer (Thermo Scientific) with 22-h runs per sample (LC as described earlier (24–26)). Briefly, peptides were loaded onto the reversed-phase strong cation exchange nano back column and desalted with a water to organic to water gradient over 15 min using 95% H_2O, 5% acetonitrile, 0.1% formic acid and 30% H_2O, 70% acetonitrile, 0.1% formic acid. The back column was then connected to a 15-cm reversed-phase resolving nano column positioned on a nanospray source (Proxeon, Thermo Scientific) directly connected to the QExactive. Peptides were eluted from the strong cation exchange column with 11 increasing salt pulses (0 to 500 mM ammonium acetate) followed by water–organic gradients (using 95% H_2O, 5% acetonitrile, 0.1% formic acid and 30% H_2O, 70% acetonitrile, 0.1% formic acid) for 2 h to resolve the peptides and subsequently ionize via nanospray into the QExactive. For all two-dimensional LC-MS/MS analyses, the QExactive was operated in data-dependent mode with the top 10 ions selected for isolation and higher-energy collisional dissociation fragmentation from the most abundant peptides from the survey scan (400–1600 *m/z*). The QExactive settings were as follows: normalized collision energy for higher-energy collisional dissociation of 28 eV, full-scan resolution of 70,000K, higher-energy collisional dissociation MS/MS resolution of 17,500, and dynamic exclusion set at 15 s. Peptides were not excluded based on charge state.

Proteome Informatics—All MS/MS spectra were searched with the SEQUEST algorithm (v. 27 (27)) and filtered with DTASelect/Contrast (v. 1.9 (28)) at the peptide level (Xcorrs of at least 1.8 (+1), 2.5 (+2), 3.5 (+3)) with a minimum DeltCN of 0.08. SEQUEST parameters included a fixed modification for carboxyamidomethylated cysteines, a variable modification for urea carbamylation of arginine and lysine residues, trypsin digestion, up to four missed cleavages, a precursor mass tolerance of 3.0 Da, and a fragment mass tolerance of 0.8 Da. Only proteins identified with two fully tryptic peptides from a 22-h run were considered for further biological study. Tandem MS/MS spectra were searched against the same *E. scolopes* protein database used for iTRAQ, but with several modifications. In addition to the squid protein sequences originating from the light organ and hemocytes, *V. fischeri* ES114 protein sequences and common contaminants (such as trypsin, keratins, and protein lab standards) were included in this combined database (22,733 protein sequences). To remove redundancy in protein sequences from the original database used during iTRAQ (34,684 sequences), protein sequences originating from hemocyte singleton transcripts were removed from the combined database. Spectral counts were extracted for all proteins and compared across all runs. False discovery rates were determined by selecting one technical run from each sym and cured proteome analysis and searching those data against a reversed decoy database created from the squid–vibrio database used during the label-free analyses as described previously (24, 29). All false positive rates were below 5% and below 0.5% when using parent peptides with a mass accuracy of $-10 <$ parts per million < 10 (Table I). Protein spectral counts were normalized and compared between sym and cured hemocytes using the normalized spectral abundance factor (NSAF) method (30, 31). Significant differences between NSAF values for proteins of sym and cured hemocytes were determined via an unpaired equal variance *t* test (p value < 0.05). For each significantly different protein, the ratio of the average NSAF for sym relative to the average NSAF for cured was Log2 transformed and filtered for proteins meeting a cutoff of a 1.5-fold change.

RESULTS

iTRAQ and label-free spectral counting were utilized to compare hemocyte proteomes from hosts with colonized or uncolonized light organs (Fig. 2). Overall, iTRAQ identified 40,771 spectra, 15,415 peptides, and 2,001 proteins at a global false discovery rate of 1% (Table I; also see the supplemental material). Of the 1,024 proteins quantified with two or more peptides, 983 had an unused protein score greater than or equal to 2.0 (99% confidence; supplemental material). General housekeeping proteins, which were not expected to change between sym and cured hemocytes (*e.g.* elongation factor 1-α (−0.04 Log2 fold change), heat shock protein 70 (−0.09 Log2 fold change), lactate dehydrogenase (0.10 Log2 fold change), and glyceraldehyde-3-phosphate dehydrogenase (0.06 Log2 fold change)), maintained a Log2 fold change near zero (supplemental material). Proteins with a Log2 fold change greater than 0.80 or less than −0.80 (corresponding to a fold change of 1.75) were considered significant after comparison of the experimental variation arising from reagent labeling and/or digestion (see "Experimental Procedures"). Of the 1,024 proteins, 37 unique proteins displayed significantly different abundances between the samples (p value < 0.05; Fig. 3). These proteins were evenly distributed between being more and less abundant. Twenty

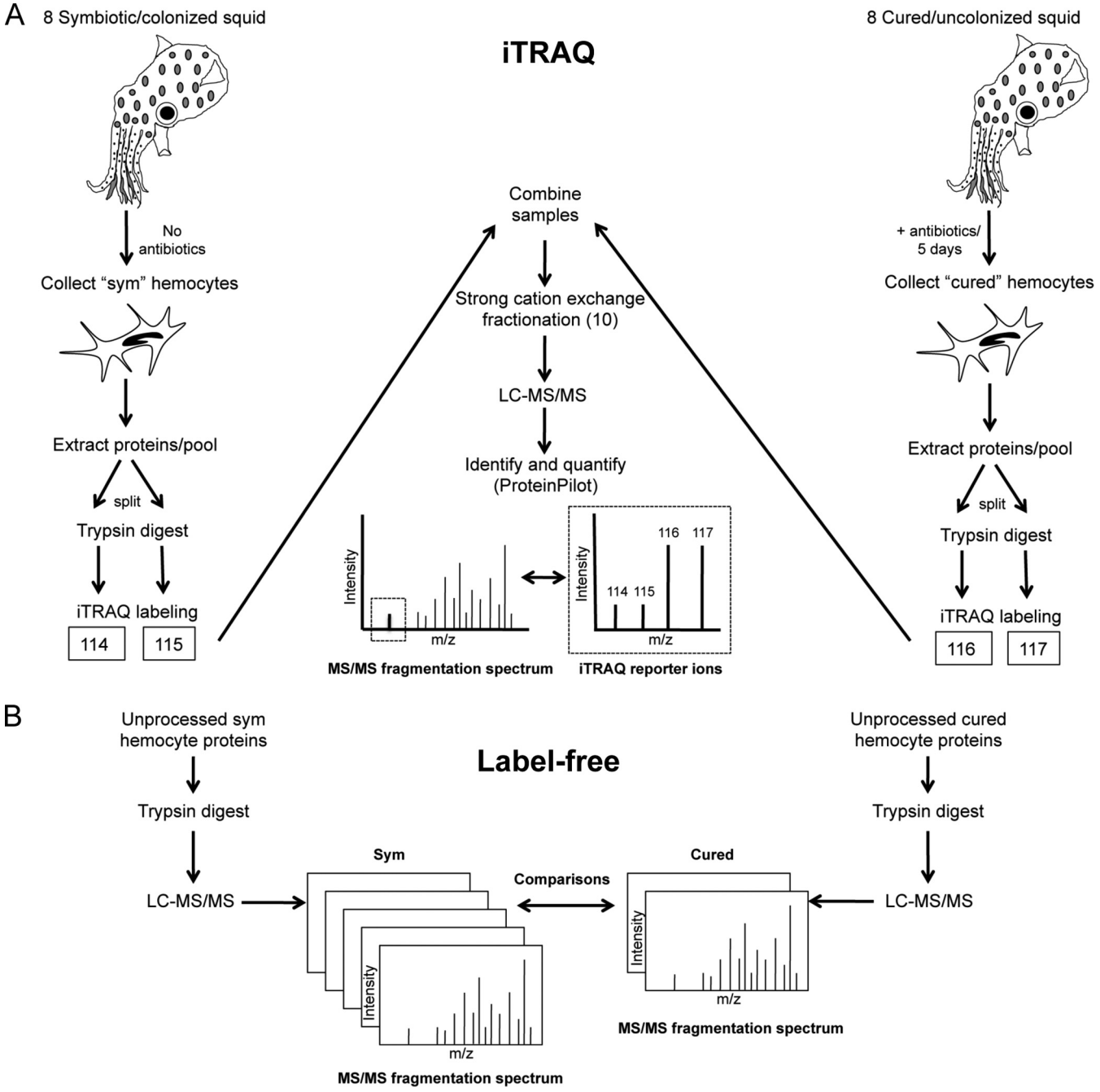

FIG. 2. **A visual diagram of the methods for collecting and quantifying hemocyte proteomes from symbiotic and cured squid using iTRAQ (*A*) and label-free proteomics (*B*) (see "Experimental Procedures").** *A*, hemocytes were collected from the cephalic blood vessels of healthy adult squid. For cured hemocytes, hemocytes were collected from squid treated with an antibiotic mixture for 5 days. Hemocytes from eight squid from each condition were pooled, and the proteins were extracted. The protein pool from sym and cured hemocytes was then split in half, resulting in experimental replicates. After digestion with trypsin, the peptides from the two sym replicates were labeled with iTRAQ reagents 114 and 115, and the two cured replicates were labeled with 116 and 117 (these reagents are isobaric mass tags that bind to the free amine groups of tryptic peptides). The four samples were combined and then fractionated into 10 fractions using strong cation exchange. LC-MS/MS was performed on each individual fraction. The spectral data from each fraction were combined and searched against an *E. scolopes* protein database using ProteinPilot. After peptide identification, the different intensities of iTRAQ reagent ions (found in the low-*m/z* region of the MS/MS spectrum) allowed the comparison of sym and cured hemocyte proteomes. The example in this figure suggests that the identified peptide from cured hemocytes (116 and 117, iTRAQ reporter ions) is more abundant than the peptide found in sym hemocytes (114 and 115). *B*, for label-free proteomics, unprocessed sym or cured hemocyte proteins from the iTRAQ experiment were digested with trypsin, subjected to LC-MS/MS, and analyzed using SEQUEST. For quantification, the normalized spectral abundance factor (NSAF) was determined for all proteins in each sample per replicate. The average NSAFs between sym and cured hemocyte proteins were compared to identify differences.

TABLE I
Summary of iTRAQ and label-free spectral counting results

Method	iTRAQ	Label-free			
Sample	Sym and cured	Sym 1	Sym 2	Cured 1	Cured 2
Spectra	40,771	69,007	69,503	35,926	40,805
Peptides	15,415	23,339	19,215	12,444	10,337
Proteins	2,001	3,313	2,923	1,713	1,506
Parts per million	N/A	83.47	81.33	80.95	81.59
Proteins quantified	1,024	1,241*			
Significant proteins	37	185			
More abundant	20	92			
Less abundant	17	93			
FDR (FDR ppm)[a]	1%	4.63% (0.48%)		2.96% (0.33%)	

[a] False discovery rate (FDR) in parentheses represents the FDR when using parent peptides with a mass accuracy of $-10 <$ parts per million (ppm) < 10; applied to the label-free approach only. * shared between all technical runs.

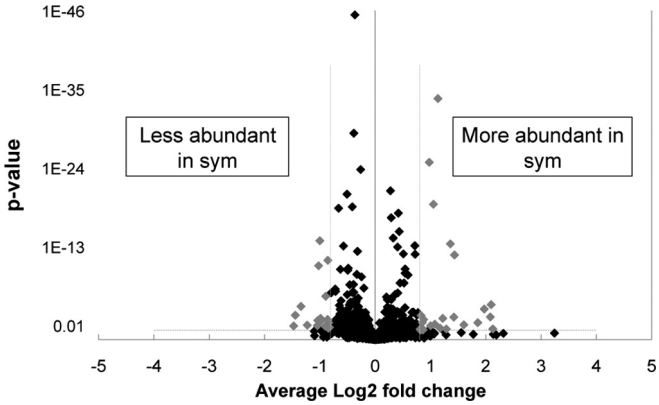

FIG. 3. **Volcano plot of hemocyte iTRAQ data.** The average Log2 fold change and respective *p* value for each protein identified using iTRAQ with more than two peptides was plotted. Proteins identified as significantly more abundant or less abundant in sym hemocytes are highlighted in gray (p value < 0.05). Gray dotted lines highlight the Log2 fold change cutoffs (± 0.80) and a p value cutoff of 0.05.

proteins were more abundant and 17 were less abundant in sym hemocytes (Table II). The protein with the greatest increase in abundance was a ganglioside GM2 activator precursor (2.12 Log2 fold change; Table II), and a hypothetical protein demonstrated the greatest reduction (-1.47 Log2 fold change; Table II). These differentially produced proteins were categorized into several functional groups, including cell adhesion, cytoskeleton, innate immunity, and lysosomal (Fig. 4).

To further validate the iTRAQ data, the same unprocessed starting samples from iTRAQ were analyzed via two-dimensional LC-MS/MS and label-free spectral counting. Each sample had a false discovery rate less than 5% (4.63% for sym and 2.96% for cured; Table I). The false discovery rates were less than 0.5% when we used peptides with a mass accuracy of $-10 <$ parts per million < 10 (Table I). NSAFs were used to determine the relative abundance of each protein identified. A total of 1,241 proteins were shared between all technical runs (supplemental material). From these analyses we found 185 unique proteins that were significantly more or less abundant in sym hemocytes (p value < 0.05; supplemental material). The distribution of significantly different sym and cured proteins showed a trend similar to what was found with iTRAQ, although more proteins were identified with the label-free technique (supplemental material). Comparison of these proteins to the iTRAQ data revealed that 33 of the 37 unique iTRAQ proteins were also identified using the label-free method (Table III). Thirty of these 33 proteins followed the same trend identified by iTRAQ (Table III). Only three of the significant iTRAQ proteins were identified via the label-free method as having an opposite trend (Table III). All three of these proteins (s-formylglutathione hydrolase, rootletin, and mps one binder kinase activator) generally had a lower number of peptides contributing to the iTRAQ ratio, which might explain the opposite trends observed with the two techniques. Although there was a large overlap between all of the iTRAQ proteins and all of the label-free proteins identified (795 in common), only 16 proteins were found to be significantly different between sym and cured hemocytes with both methods (p value < 0.05, Table III; Fig. 5). In comparison to the iTRAQ data, the label-free analysis identified an additional 169 unique proteins (Fig. 5). Among the significantly more abundant proteins identified by both methods were cathepsin L2 and a ganglioside GM2 activator precursor (Table III). The less abundant proteins mutually identified by iTRAQ and spectral counting included myosin regulatory light chain 2, a LIM domain and actin binding protein, and a surfactant B-like protein (Table III). One common limitation to iTRAQ is ratio suppression, which can lead to smaller fold changes and therefore a lesser number of significantly different proteins detected (32, 33). However, analysis of the significant proteins identified by label-free spectral counting revealed categories of proteins similar to the proteins identified by iTRAQ (Fig. 6). Proteins associated with the cytoskeleton and cell adhesion were generally less abundant (16 of 20), whereas proteases were mostly more abundant (4 of 5) in sym hemocytes. Other differences were also detected for proteins associated with innate immunity, lysosomes, and stress responses (Fig. 6).

An antibody generated to recognize a peptide specific to cathepsin L2, a protein that demonstrated one of the more significant increases in sym hemocytes in both iTRAQ and spectral counting analyses, was used to compare the local-

TABLE II
Proteins identified by iTRAQ as significantly more or less abundant in sym hemocytes

Trend[a]	Database I.D.	Protein	Pep[b]	Log2 fold change[c]	*p* value
More abundant	Meta990274633	Cathepsin D	39	1.054	1.20E-19
	Meta990279877	Cathepsin L2	12	2.097	1.29E-05
	Meta990255245	Galectin 4-like protein transcript variant	2	1.287	3.21E-02
	Meta990274730	Ganglioside GM2 activator precursor	4	1.222	1.08E-03
	Meta990237061	Ganglioside GM2 activator precursor	6	2.082	6.87E-04
	Meta990280234	Ganglioside GM2 activator precursor	4	2.128	3.18E-02
	Meta990084014	Glucosamine-fructose-6-phosphate aminotransferase	2	0.820	4.50E-02
	Meta990256250	Histone H4-I	56	1.135	1.99E-34
	Meta990272817	26S proteasome non-ATPase subunit 12	2	0.825	4.26E-02
	Meta990275393	Hypothetical protein BRAFLDRAFT 185672	5	0.829	4.33E-04
	Meta990279786	Pancreatic triacylglycerol lipase-like	4	1.856	4.30E-03
	Meta990297914	Amino acid transport protein rBAT	8	0.877	1.58E-03
	Meta990261223	Pyruvate carboxylase	2	0.830	3.02E-02
	Meta990234931	Histone H3.2-like	24	1.434	1.46E-12
	Meta990123568	Intermediate filament protein	5	0.853	3.93E-03
	Meta990236123	Intermediate filament protein	12	1.972	5.32E-05
	Meta990244043	l-3-hydroxyacyl-coenzyme a dehydrogenase	4	0.862	4.43E-02
	Meta990233579	LeechCAM	5	1.422	7.15E-04
	Meta990279781	Myosin heavy chain	4	0.869	3.98E-02
	Meta990231669	Myosin heavy chain isoform B	66	0.978	1.64E-25
	Meta990231856	Aconitate hydratase, mitochondrial-like	4	1.078	6.88E-03
	Meta990275031	s-formylglutathione hydrolase-like	5	0.812	3.57E-02
	Meta990210359	WW domain-containing oxidoreductase-like	20	1.359	4.13E-14
	Meta990236159	WW domain-containing oxidoreductase-like	4	1.603	8.54E-03
	Meta990272315	Rab3	3	1.176	2.51E-02
Less abundant	Meta990253072	GM11222	5	−0.844	2.56E-02
	Meta990275900	Heat shock protein 22 isoform 1	7	−0.975	2.13E-03
	Meta990300071	Hypothetical protein	3	−0.858	1.51E-03
	Meta990081459	Hypothetical protein	8	−1.341	2.30E-05
	Meta990232278	LIM domain and actin-binding protein 1	29	−1.000	1.52E-14
	Meta990297867	Myosin regulatory light chain 2	26	−0.856	8.14E-12
	Meta990275793	EsPGRP 5	25	−1.023	4.71E-11
	Meta990299707	Calumenin-like	3	−1.227	9.64E-03
	Meta990236998	Calumenin-like	6	−0.988	1.08E-03
	Meta990232620	Dynactin 2-like	6	−1.045	1.13E-02
	Meta990220758	Hypothetical protein LOC100115922	3	−1.474	1.23E-02
	Meta990232946	mps one binder kinase activator-like 2B	3	−0.807	2.46E-02
	Meta990231811	Rootletin-like	2	−0.904	1.35E-02
	Meta990232518	Protein kinase, AMP-activated, β 1	5	−0.839	4.23E-03
	Meta990274714	Similar to surfactant B protein	8	−1.037	2.10E-03
	Meta990255307	THO complex 4-like	3	−1.444	3.91E-04
	Meta990299300	Translationally controlled tumor protein	18	−0.894	8.92E-07
	Meta990139263	Tropomyosin	4	−0.826	1.89E-02

[a] Protein trend based on iTRAQ data.
[b] Pep: number of peptides contributing to the iTRAQ ratio for that protein.
[c] Log2 fold change represents a protein difference in sym relative to cured.

ization and abundance of this cathepsin in sym and cured hemocytes by immunocytochemistry. Fluorescence measurements of composite confocal images of individual hemocytes revealed that cathepsin L2 was significantly more abundant in sym hemocytes (+1.79-fold), providing further confirmation of the mass-spectrometry-based proteomic results (Fig. 7). In addition, the localization of cathepsin L2 in sym hemocytes appeared to be organized in punctate granules consistent with lysosomes (Fig. 7*A*; for negative controls, see the supplemental material). In contrast, the localization of cathepsin L2 in cured hemocytes appeared to be more diffuse and present both inside and outside of granules (Fig. 7*A*).

DISCUSSION

The *E. scolopes*–*V. fischeri* association presents a unique opportunity to study the influence of the innate immune system on beneficial symbiosis (2, 8). Because there is a high degree of specificity such that only *V. fischeri* can colonize the light organ, mechanisms must be in place to safeguard the

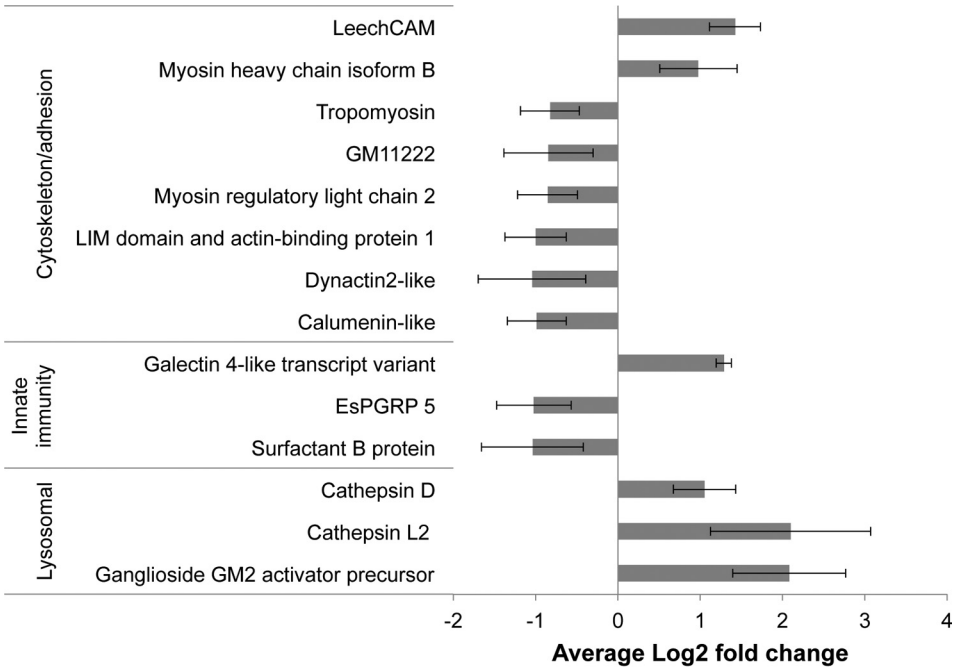

FIG. 4. **Proteins from sym hemocytes that were significantly more or less abundant (via iTRAQ).** Proteins with the most peptides and the largest Log2 fold changes were selected to represent a group of similar proteins (Table II).

association. Host hemocytes are believed to play a significant role in this process, and a previous study revealed that these cells are able to differentiate between *V. fischeri* and non-symbiotic bacteria, even in the absence of an adaptive immune system (3). The molecular mechanisms that allow these hemocytes to make this distinction are poorly understood, but some target genes and proteins have been identified (10). In this study, the hemocyte proteome from *E. scolopes* was characterized based on colonization of the light organ in order to identify host proteins that may be involved with symbiont recognition. Quantitative proteomics is becoming a reliable technique for characterizing host–microbe interactions during symbiosis, and both label and label-free methods have been applied to beneficial associations involving plants, insects, corals, and humans (26, 34–38). This study represents, to the best of our knowledge, the first use of two complementary quantitative proteomic approaches to further understand the influence of a bacterial symbiont on the immune response of an animal host.

Cytoskeletal Dynamics and Cell Adhesion—Many of the hemocyte proteins influenced by *V. fischeri* colonization were related to the cytoskeleton. Both iTRAQ and spectral counting supported this observation; however, the label-free approach resulted in the identification of a greater number of differentially produced cytoskeletal and adhesion proteins (iTRAQ, 10 of 37; label-free, 20 of 185; see Table II, Fig. 6, supplemental material). Proteins associated with integrin signaling were also identified with both methods (Table II, Fig. 6). Although involved in a variety of cellular pathways, integrin signaling is a well-characterized component of cell adhesion and cytoskeletal rearrangements in other types of macrophages (39, 40). The less abundant cytoskeletal/adhesion proteins identified in this study might represent a decrease in the ability of the hemocytes to adhere to and/or phagocytose *V. fischeri*. Previous results have shown that *V. fischeri* binds significantly less to hemocytes from a colonized host than to cells from one that has been cured (3). Bacteria can influence the host cell cytoskeleton in a number of ways, especially as an effective strategy for immune evasion (41, 42). In the context of the squid–vibrio symbiosis, it has been shown that *V. fischeri* induces changes to the host's actin cytoskeleton in the tissues surrounding the light organ (43). The mechanism by which *V. fischeri* may deliver a signal that alters the hemocyte cytoskeleton remains unknown. However, the outer membrane protein OmpU appears to be involved with mediating adhesion to host hemocytes in both the squid–vibrio symbiosis and the pathogenic association between oysters and *Vibrio splendidus* (3, 44, 45). The exact mechanism by which OmpU or some other bacterial factor(s) mediates the hemocyte response remains to be characterized, but the results of this study suggest that colonization by *V. fischeri* does alter the hemocyte cytoskeleton.

Other hemocyte proteins putatively associated with cell adhesion and innate immunity were a leech cell adhesion molecule, surfactant-like proteins, and a galectin 4–like protein (Tables II and III; Fig. 6). In leech neurons, cell adhesion molecule is a transmembrane protein that has been implicated in cell attachment and migration (46–49). Two putative

TABLE III
Comparison of proteins identified by iTRAQ with NSAF values from the label-free analysis

iTRAQ[a]	Database I.D.	Protein	NSAF sym[b]	NSAF cured[b]	*p* value[c]	LF[d]
More abundant	Meta990274633	Cathepsin D	1.42E-03	1.24E-03	0.388	↑
	Meta990279877	**Cathepsin L2**	**1.83E-03**	**8.50E-05**	**0.016**	↑
	Meta990255245	Galectin 4-like protein transcript variant	2.20E-04	4.00E-05	–	↑
	Meta990274730	**Ganglioside GM2 activator precursor**	**4.50E-04**	**2.70E-04**	**0.038**	↑
	Meta990237061	Ganglioside GM2 activator precursor	2.80E-04	N/a	–	↑
	Meta990280234	Ganglioside GM2 activator precursor	7.50E-04	N/a	–	↑
	Meta990084014	Glucosamine-fructose-6-phosphate aminotransferase	N/a	N/a	–	N/a
	Meta990256250	Histone H4-I	2.70E-03	2.52E-03	0.683	↑
	Meta990272817	26S proteasome non-ATPase subunit 12	3.85E-04	1.65E-04	0.174	↑
	Meta990275393	**Hypothetical protein BRAFLDRAFT 185672**	**6.45E-04**	**2.05E-04**	**0.011**	↑
	Meta990279786	**Pancreatic triacylglycerol lipase-like**	**2.37E-03**	**7.30E-04**	**0.006**	↑
	Meta990297914	**Amino acid transport protein rBAT**	**6.35E-04**	**1.45E-04**	**0.020**	↑
	Meta990261223	Pyruvate carboxylase	2.40E-04	N/a	–	↑
	Meta990234931	Histone H3.2-like	1.57E-03	7.05E-04	0.202	↑
	Meta990123568	Intermediate filament protein	N/a	N/a	–	N/a
	Meta990236123	**Intermediate filament protein**	**8.80E-04**	**7.00E-05**	**0.010**	↑
	Meta990244043	l-3-hydroxyacyl-coenzyme a dehydrogenase	5.80E-04	1.85E-04	0.074	↑
	Meta990233579	LeechCAM	1.80E-04	4.50E-05	0.123	↑
	Meta990279781	**Myosin heavy chain**	**3.40E-04**	**1.10E-04**	**0.044**	↑
	Meta990231669	Myosin heavy chain isoform B	7.05E-04	4.20E-04	0.062	↑
	Meta990231856	**Aconitate hydratase, mitochondrial-like**	**3.15E-04**	**1.15E-04**	**0.016**	↑
	Meta990275031	s-formylglutathione hydrolase-like	2.60E-04	3.15E-04	0.409	↓
	Meta990210359	WW domain-containing oxidoreductase-like	N/a	N/a	–	N/a
	Meta990236159	**WW domain-containing oxidoreductase-like**	**1.08E-03**	**3.25E-04**	**0.017**	↑
	Meta990272315	Rab3	5.95E-04	3.10E-04	0.224	↑
Less abundant	Meta990253072	GM11222	5.45E-04	7.35E-04	0.127	↓
	Meta990275900	**Heat shock protein 22 isoform 1**	**4.00E-05**	**3.35E-04**	**0.033**	↓
	Meta990300071	Hypothetical protein	N/a	N/a	–	N/a
	Meta990081459	Hypothetical protein	N/a	N/a	–	N/a
	Meta990232278	**LIM domain and actin-binding protein 1**	**7.50E-04**	**1.73E-03**	**0.012**	↓
	Meta990297867	**Myosin regulatory light chain 2**	**2.12E-03**	**5.12E-03**	**0.007**	↓
	Meta990275793	EsPGRP 5	1.85E-04	3.25E-04	0.067	↓
	Meta990299707	Calumenin-like	2.15E-04	8.75E-04	0.158	↓
	Meta990236998	Calumenin-like	4.00E-04	1.12E-03	0.235	↓
	Meta990232620	**Dynactin 2-like**	**1.90E-04**	**6.05E-04**	**0.010**	↓
	Meta990220758	**Hypothetical protein LOC100115922**	**2.95E-04**	**7.20E-04**	**0.010**	↓
	Meta990232946	mps one binder kinase activator-like 2B-like	4.00E-05	1.00E-04	–	↑
	Meta990231811	Rootletin-like	5.00E-05	3.00E-05	–	↑
	Meta990232518	Protein kinase, AMP-activated, β 1	3.83E-03	5.97E-03	0.161	↓
	Meta990274714	**Similar to surfactant B protein**	**9.00E-05**	**3.25E-04**	**0.002**	↓
	Meta990255307	**THO complex 4-like**	**1.00E-04**	**1.80E-04**	**0.015**	↓
	Meta990299300	Translationally controlled tumor protein	1.55E-03	3.36E-03	0.079	↓
	Meta990139263	Tropomyosin	N/a	N/a	–	N/a

[a] Trend represents sym hemocytes relative to cured hemocytes.

[b] Normalized spectral abundance factor (NSAF). NSAFs were calculated as an average of two technical replicates per hemocyte condition.

[c] Only proteins identified in both replicates of sym and cured hemocytes were used to calculate a *p* value. Proteins marked "N/A" were not identified using the label-free approach. Trends were still determined if a *p* value was not calculated. Proteins in bold were identified as significant by both iTRAQ and spectral counting.

[d] Label-free (LF) trend as determined by comparing the average NSAF value of the specific sym hemocyte protein relative to the same cured hemocyte protein. Up arrows indicate proteins that were more abundant, and down arrows indicate proteins that were less abundant.

saposin/surfactant-like proteins were identified as less abundant via both methods. In other systems these proteins have the ability to bind microorganisms and enhance phagocytosis (50, 51). Future studies will focus on determining whether these surfactant-like proteins in *E. scolopes* are involved with the observed increased binding of *V. fischeri* to cured hemocytes. Additionally, iTRAQ and spectral counting identified one galectin (a galectin 4–like transcript variant) as more abundant in sym hemocytes. Galectins are carbohydrate-binding proteins that are believed to recognize the surface-exposed carbohydrates of microorganisms (52, 53). Many galectins are now implicated as important factors of the invertebrate innate immune response, as these proteins can directly bind bacteria and enhance phagocytosis (54, 55).

Recently in the lancelet *Branchiostoma belcheri tsingtauense*, it was discovered that a galectin bound to *Vibrio vulnificus*, but not to *Vibrio parahemeolyticus* (56). The apparent selectivity of galectins might also contribute to the ability of *E. scolopes* to differentiate between closely related *Vibrio* species, but this remains to be characterized.

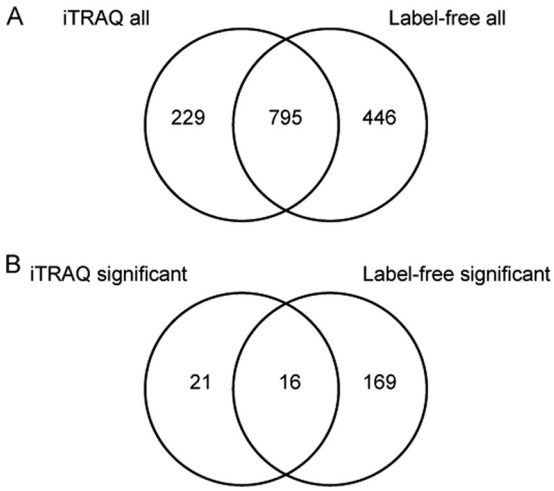

FIG. 5. **Comparison of iTRAQ and label-free proteomic analyses of sym and cured hemocytes.** *A*, Venn diagram showing the overlap of all iTRAQ proteins identified (>2 peptides) and all label-free proteins identified (>2 peptides, identified in both samples and in both replicates). *B*, Venn diagram of the differentially produced proteins identified using iTRAQ and label-free methods (16 proteins were significant in both iTRAQ and label-free analyses).

Taken together, the data from this and previous studies suggest that a specific response to *V. fischeri* may require coordination among host proteins involved with microbial recognition, adhesion, and engulfment. It is unclear how a putative decrease in cytoskeletal function might be involved with a specific response to *V. fischeri*, as sym and cured hemocytes still have the ability to adhere to and phagocytose non-symbiotic bacteria (3). As an alternative hypothesis, the change in abundance of cell adhesion and cytoskeletal proteins might influence hemocyte migration. The light organ is highly vascularized, and hemocytes are known to migrate into the light organ crypt spaces where the symbiont is found at a high density (3, 11, 12). Future studies should determine whether hemocytes from sym and cured hosts migrate to the light organ with the same frequency and numbers.

Lysosomal Proteins—Some of the more significantly abundant proteins are predicted to be involved with lysosome function. Similar to other macrophage-like cells, hemocytes of *E. scolopes* have lysosomes that play an important role in the immune response of phagocytic cells by degrading engulfed bacteria (12, 57). Cathepsins and a ganglioside GM2 activator

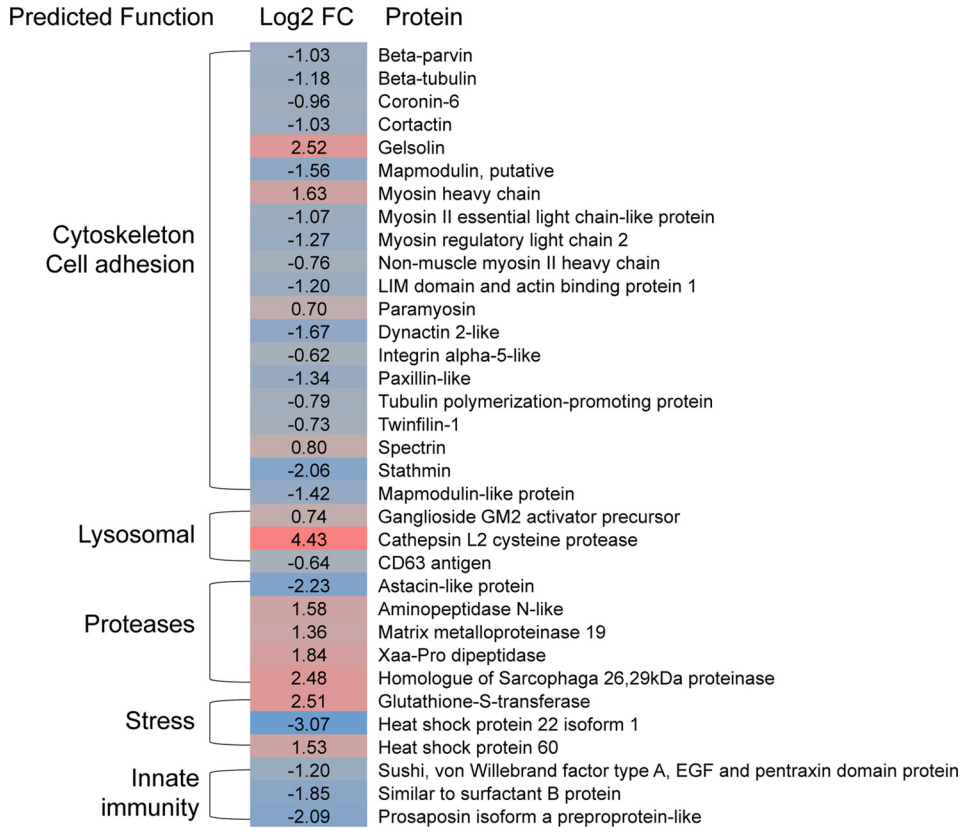

FIG. 6. **Heat map of proteins significantly more or less abundant in sym hemocytes identified via label-free spectral counting.** For each protein, an average NSAF ratio of sym relative to cured hemocytes was Log2 transformed. Proteins in blue were less abundant, and those in red were more abundant. The scale ranges from a −3.1 Log2 fold change (blue) to a 4.5 Log2 fold change (red).

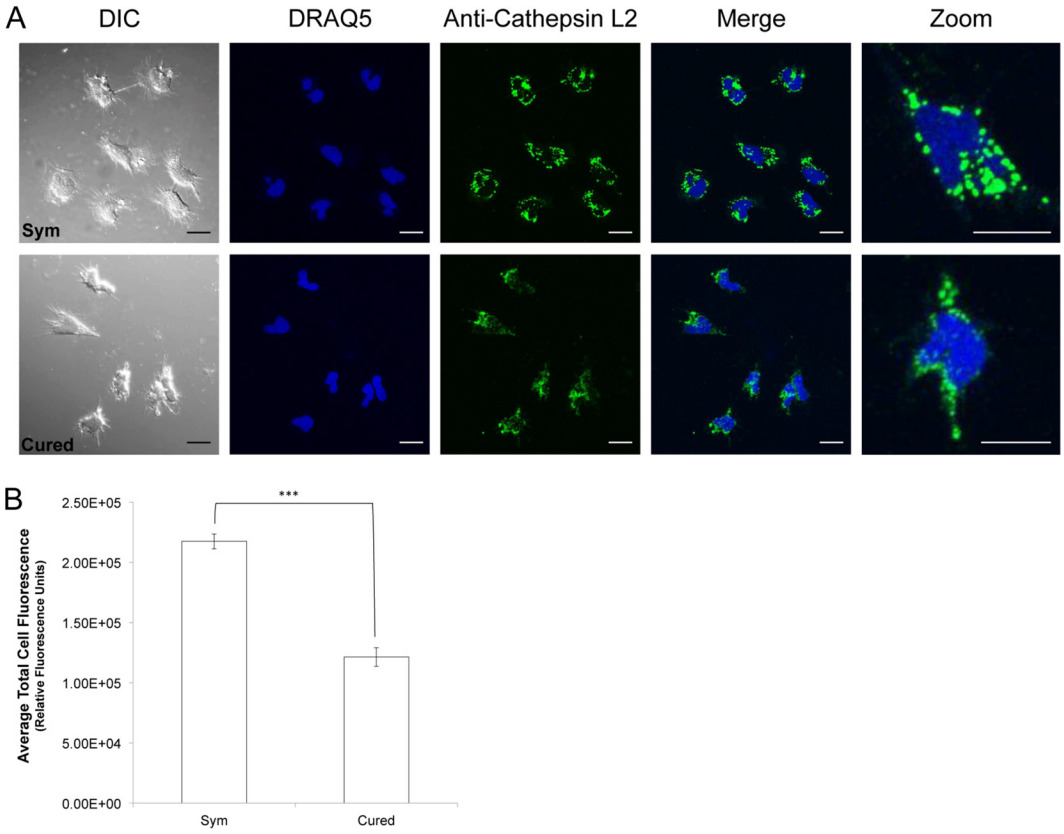

FIG. 7. **Protein localization of cathepsin L2 in sym and cured hemocytes.** *A*, representative maximum projection composite confocal images of sym and cured hemocytes. Anti-cathepsin L2 was detected with an Alexa Fluor 488 goat anti-rabbit secondary antibody. The hemocytes were counterstained with the nuclear stain DRAQ5. A higher-magnification view of additional representative hemocytes revealed that cathepsin L2 was abundant and localized in punctate granules consistent with lysosomes in sym hemocytes. In hemocytes from cured animals, cathepsin L2 was less intense and had a more diffuse staining pattern (DIC, differential interference contrast image; scale bar measurements = 10 μm). Hemocyte negative controls (secondary antibody only) can be found in the supplemental material. *B*, fluorescence measurements from maximum projections of individual hemocytes were utilized to compare the abundance of cathepsin L2 in sym (n = 3 squid, 89 hemocytes total) and cured hemocytes (n = 2 squid, 28 hemocytes total). Normalized cell fluorescence was determined with the software FIJI by measuring the integrated density of the cell and then subtracting the mean background fluorescence multiplied by the area of the cell measured. Cathepsin L2 was significantly more abundant in sym hemocytes (1.79-fold; ***two-tailed unequal variance t test, p value < 0.0001; error bars represent standard error of the mean (S.E.)).

were more abundant in sym hemocytes (Tables II and III). Although linked to other cellular pathways, cathepsins are proteases activated by a low lysosomal pH, and the ganglioside GM2 activator precursor is a lipid transfer protein also associated with lysosomes (58, 59). As part of normal lysosomal function, these proteins may be important for processing microbe-associated molecular patterns from phagocytosed bacteria. The differences found in lysosomal and other proteins identified in this study might not be restricted to hemocytes. For example, a recent study analyzed the influence of *V. fischeri* on the host's light organ transcriptome during the initiation of the association and found that both cathepsin D and L2, along with the GM2 activator precursor protein and a triacylglycerol lipase, were also significantly more abundant in symbiotic light organs (21). Further examination using an antibody generated to cathepsin L2 corroborated that colonization by *V. fischeri* increases the abundance of this protease in hemocytes (Fig. 7). Cathepsin L–like proteins have been described in other invertebrate hemocytes (60, 61) and are modulated by colonization in both pathogenic (62, 63) and beneficial associations (Ref. 64 and this study).

Additional Innate Immunity Proteins—A transcriptome and proteome of circulating hemocytes from symbiotic hosts revealed a number of genes and proteins related to innate immunity, including members of the NF-κB pathway, pattern recognition receptors (two PGRPs and a galectin), and complement-like proteins (10). Building on this previous research, iTRAQ revealed that the EsPGRP 5 protein is less abundant in sym hemocytes, although transcript levels of this gene were up-regulated (Table II, Fig. 4) (10). The retention time and turnover of EsPGRP 5 are unknown but might explain this discrepancy. EsPGRP 5 also has a putative signal peptide, and secretion of this protein might be influenced by the presence or absence of *V. fischeri* (10). The cellular localizations of two other PGRPs from *E. scolopes* (EsPGRP 1 and EsPGRP 2) are also altered in response to colonization of the host light

organ (10, 65, 66). iTRAQ does not quantify extracellular proteins lost during sample preparation, so colonization might lead to an increased release of EsPGRP 5 and therefore a loss of the cellular protein, although this remains to be tested.

A matrix metalloproteinase, a protein also associated with innate immunity, was significantly more abundant in sym hemocytes analyzed via label-free spectral counting (Fig. 6). This matrix metalloproteinase with homology to a complement-related protein, vitronectin, was previously identified by means of transcriptomics and shotgun proteomics (10). Unlike for EsPGRP 5, an increase in protein abundance for matrix metalloproteinase did correlate with previously identified increases in gene expression (10). Recent in-depth comparisons of mRNA and protein levels in yeast, mice, and humans have demonstrated that mRNA abundance does not always correlate with protein abundance. In these studies ~40% of mRNA transcripts predicted protein abundance (67, 68). The variation between the rest of the transcripts and proteins is likely due to the complex dynamics associated with regulating post-transcriptional, post-translational, and protein degradation processes. These pathways are poorly understood in the squid–vibrio symbiosis, but a global transcriptomic comparison between sym and cured hemocytes is currently underway and will be useful for future analyses.

Summary—The mechanisms by which hemocytes of *E. scolopes* differentiate between the light organ symbiont *V. fischeri* and non-symbiotic bacteria are still under investigation, but it is clear that colonization influences the cellular response, gene expression, and proteome of this cell type. Previous proteomic analyses of the squid–vibrio symbiosis have been focused on two-dimensional gel electrophoresis or individual protein identification and characterization (10, 13, 69). This study represents the first application of quantitative proteomics to the squid–vibrio symbiosis. In addition, this study represents the first combination of iTRAQ and a label-free proteomic analysis to enhance our understanding of a beneficial symbiosis. Proteomic and gene expression studies in other organisms, including oysters and pigs, have revealed that similar genes and proteins (cathepsins, GM2AP, heat shock protein 22, and cytoskeletal proteins) are altered by both pathogenic and beneficial bacterial colonization (70, 71). Few studies have combined both iTRAQ and label-free proteomics; however, label-free methods have been shown to identify a greater number of significantly different proteins than labeling approaches, similar to what was observed in this study (72, 73). Employing both methods, however, might offer a means to identify those proteins or pathways displaying the most significant changes.

In this study, after curing, both quantitative proteomic techniques (iTRAQ and spectral counting) revealed significant differences between hemocytes from colonized and uncolonized squid, bolstering the hypothesis that *V. fischeri* influences host hemocyte function. Although the mechanisms are still unclear, the proteins identified independently via both techniques suggest that a number of cellular processes (*e.g.* cytoskeletal, lysosomal, and immune recognition) may help mediate host tolerance of *V. fischeri*. The mechanisms by which invertebrate innate immune systems target and recognize specific microorganisms are poorly understood, and the results from this study might have broader implications for understanding these processes in other taxa. This study acts as a foundation for understanding how the hemocyte proteome changes in response to colonization, and future studies will focus on characterizing hemocyte proteins with respect to their role in the *E. scolopes–V. fischeri* symbiosis. In addition, directly challenging sym and cured hemocytes with symbiotic and non-symbiotic bacteria will likely help identify the dynamics involved with host recognition of bacteria.

Acknowledgments—T. Abbott, M. LoPresti, and C. Colangelo of the W.M. Keck Mass Spectrometry and Protein Chemistry Facility at Yale University provided assistance with the iTRAQ analyses. We thank B. Rader for assistance with squid collections and the Kewalo Marine Laboratory of the University of Hawaii for providing access to research facilities. P. Lapierre, M. McFall-Ngai, and E. Ruby assisted with generating the squid protein sequence database. We also thank M. McFall-Ngai (University of Wisconsin) for the donation of the cathepsin antibody, which was funded by NIH NI50661 to M. McFall-Ngai and E. G. Ruby. We thank C. Norris of the University of Connecticut Flow Cytometry and Confocal Microscopy Facility for assistance with the immunocytochemistry experiments. Finally, we thank A. Kerwin and A. Collins for their comments on the manuscript.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [74, 75] with the dataset identifier PXD001185.

* This work was supported by the National Science Foundation IOS-0958006, the University of Connecticut Research Foundation to S.V.N., and the Antonio H. and Marjorie J. Romano Graduate Education Fellowship to T.R.S.

S This article contains supplemental material.

** To whom correspondence should be addressed: Spencer V. Nyholm, Department of Molecular and Cell Biology, University of Connecticut, 91 N. Eagleville Rd., Unit 3125, Storrs, CT 06269. E-mail: spencer.nyholm@uconn.edu.

¶ Current address: Berg Diagnostics (Division of BERG LLC), Framingham, MA 01701.

REFERENCES

1. McFall-Ngai, M. (2007) Adaptive immunity: care for the community. *Nature* **445,** 153
2. Nyholm, S. V., and Graf, J. (2012) Knowing your friends: invertebrate innate immunity fosters beneficial bacterial symbioses. *Nat. Rev. Microbiol.* **10,** 815–827
3. Nyholm, S. V., Stewart, J. J., Ruby, E. G., and McFall-Ngai, M. J. (2009) Recognition between symbiotic *Vibrio fischeri* and the haemocytes of *Euprymna scolopes. Environ. Microbiol.* **11,** 483–493
4. Canesi, L., Gallo, G., Gavioli, M., and Pruzzo, C. (2002) Bacteria-hemocyte interactions and phagocytosis in marine bivalves. *Microsc. Res. Tech.* **57,** 469–476
5. Marmaras, V. J., and Lampropoulou, M. (2009) Regulators and signalling in insect haemocyte immunity. *Cell Signal.* **21,** 186–195
6. Pham, L. N., Dionne, M. S., Shirasu-Hiza, M., and Schneider, D. S. (2007) A specific primed immune response in *Drosophila* is dependent on phagocytes. *PLoS Pathog.* **3,** e26
7. Rodrigues, J., Brayner, F. A., Alves, L. C., Dixit, R., and Barillas-Mury, C. (2010) Hemocyte differentiation mediates innate immune memory in

Anopheles gambiae mosquitoes. *Science* **329,** 1353–1355
8. McFall-Ngai, M., Nyholm, S. V., and Castillo, M. G. (2010) The role of the immune system in the initiation and persistence of the *Euprymna scolopes–Vibrio fischeri* symbiosis. *Semin. Immunol.* **22,** 48–53
9. Nyholm, S. V., and McFall-Ngai, M. J. (2004) The winnowing: establishing the squid-vibrio symbiosis. *Nat. Rev. Microbiol.* **2,** 632–642
10. Collins, A. J., Schleicher, T. R., Rader, B. A., and Nyholm, S. V. (2012) Understanding the role of host hemocytes in a squid/vibrio symbiosis using transcriptomics and proteomics. *Front. Immunol.* **3,** 91
11. Koropatnick, T. A., Kimbell, J. R., and McFall-Ngai, M. J. (2007) Responses of host hemocytes during the initiation of the squid-Vibrio symbiosis. *Biol. Bull.* **212,** 29–39
12. Nyholm, S. V., and McFall-Ngai, M. J. (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
13. Schleicher, T. R., and Nyholm, S. V. (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
14. Collins, A. J., and Nyholm, S. V. (2010) Obtaining hemocytes from the Hawaiian bobtail squid *Euprymna scolopes* and observing their adherence to symbiotic and non-symbiotic bacteria. *J. Vis. Exp.* **36:** 1714
15. Heath-Heckman, E. A., and McFall-Ngai, M. J. (2011) The occurrence of chitin in the hemocytes of invertebrates. *Zoology* **114,** 191–198
16. Burgess, A., Vigneron, S., Brioudes, E., Labbe, J. C., Lorca, T., and Castro, A. (2010) Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 12564–12569
17. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682
18. Wessel, D., and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138,** 141–143
19. Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M., and Schaeffer, D. A. (2007) The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol. Cell. Proteomics* **6,** 1638–1655
20. Chun, C. K., Scheetz, T. E., Bonaldo Mde, F., Brown, B., Clemens, A., Crookes-Goodson, W. J., Crouch, K., DeMartini, T., Eyestone, M., Goodson, M. S., Janssens, B., Kimbell, J. L., Koropatnick, T. A., Kucaba, T., Smith, C., Stewart, J. J., Tong, D., Troll, J. V., Webster, S., Winhall-Rice, J., Yap, C., Casavant, T. L., McFall-Ngai, M. J., and Soares, M. B. (2006) An annotated cDNA library of juvenile *Euprymna scolopes* with and without colonization by the symbiont *Vibrio fischeri. BMC Genomics* **7,** 154
21. Kremer, N., Philipp, E. E., Carpentier, M. C., Brennan, C. A., Kraemer, L., Altura, M. A., Augustin, R., Hasler, R., Heath-Heckman, E. A., Peyer, S. M., Schwartzman, J., Rader, B. A., Ruby, E. G., Rosenstiel, P., and McFall-Ngai, M. J. (2013) Initial symbiont contact orchestrates host-organ-wide transcriptional changes that prime tissue colonization. *Cell Host Microbe* **14,** 183–194
22. Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193
23. McDonald, W. H., Ohi, R., Miyamoto, D. T., Mitchison, T. J., and Yates, J. R., III (2002) Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int. J. Mass Spectrom.* **219,** 245–251
24. Lo, I., Denef, V. J., Verberkmoes, N. C., Shah, M. B., Goltsman, D., DiBartolo, G., Tyson, G. W., Allen, E. E., Ram, R. J., Detter, J. C., Richardson, P., Thelen, M. P., Hettich, R. L., and Banfield, J. F. (2007) Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. *Nature* **446,** 537–541
25. Ram, R. J., Verberkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C., 2nd, Shah, M., Hettich, R. L., and Banfield, J. F. (2005) Community proteomics of a natural microbial biofilm. *Science* **308,** 1915–1920
26. Verberkmoes, N. C., Russell, A. L., Shah, M., Godzik, A., Rosenquist, M., Halfvarson, J., Lefsrud, M. G., Apajalahti, J., Tysk, C., Hettich, R. L., and Jansson, J. K. (2009) Shotgun metaproteomics of the human distal gut microbiota. *ISME J.* **3,** 179–189
27. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989
28. Tabb, D. L., McDonald, W. H., and Yates, J. R., 3rd (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **1,** 21–26
29. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2,** 43–50
30. Florens, L., Carozza, M. J., Swanson, S. K., Fournier, M., Coleman, M. K., Workman, J. L., and Washburn, M. P. (2006) Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40,** 303–311
31. Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L., and Washburn, M. P. (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae. J. Proteome Res.* **5,** 2339–2347
32. Karp, N. A., Huber, W., Sadowski, P. G., Charles, P. D., Hester, S. V., and Lilley, K. S. (2010) Addressing accuracy and precision issues in iTRAQ quantitation. *Mol. Cell. Proteomics* **9,** 1885–1897
33. Ow, S. Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P. C. (2009) iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly." *J. Proteome Res.* **8,** 5347–5355
34. Fan, Y., Thompson, J. W., Dubois, L. G., Moseley, M. A., and Wernegreen, J. J. (2013) Proteomic analysis of an unculturable bacterial endosymbiont (*Blochmannia*) reveals high abundance of chaperonins and biosynthetic enzymes. *J. Proteome Res.* **12,** 704–718
35. Li, X., LeBlanc, J., Truong, A., Vuthoori, R., Chen, S. S., Lustgarten, J. L., Roth, B., Allard, J., Ippoliti, A., Presley, L. L., Borneman, J., Bigbee, W. L., Gopalakrishnan, V., Graeber, T. G., Elashoff, D., Braun, J., and Goodglick, L. (2011) A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS One* **6,** e26542
36. Poliakov, A., Russell, C. W., Ponnala, L., Hoops, H. J., Sun, Q., Douglas, A. E., and van Wijk, K. J. (2011) Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. *Mol. Cell. Proteomics* **10,** M110.007039
37. Tolin, S., Arrigoni, G., Moscatiello, R., Masi, A., Navazio, L., Sablok, G., and Squartini, A. (2013) Quantitative analysis of the naringenin-inducible proteome in *Rhizobium leguminosarum* by isobaric tagging and mass spectrometry. *Proteomics* **13,** 1961–1972
38. Weston, A. J., Dunlap, W. C., Shick, J. M., Klueter, A., Iglic, K., Vukelic, A., Starcevic, A., Ward, M., Wells, M. L., Trick, C. G., and Long, P. F. (2012) A profile of an endosymbiont-enriched fraction of the coral *Stylophora pistillata* reveals proteins relevant to microbial-host interactions. *Mol. Cell. Proteomics* **11,** M111.015487
39. Berton, G., and Lowell, C. A. (1999) Integrin signalling in neutrophils and macrophages. *Cell Signal.* **11,** 621–635
40. Dupuy, A. G., and Caron, E. (2008) Integrin-dependent phagocytosis: spreading from microadhesion to new concepts. *J. Cell Sci.* **121,** 1773–1783
41. Baxt, L. A., Garza-Mayers, A. C., and Goldberg, M. B. (2013) Bacterial subversion of host innate immune pathways. *Science* **340,** 697–701
42. Ham, H., Sreelatha, A., and Orth, K. (2011) Manipulation of host membranes by bacterial effectors. *Nat. Rev. Microbiol.* **9,** 635–646
43. Kimbell, J. R., and McFall-Ngai, M. J. (2004) Symbiont-induced changes in host actin during the onset of a beneficial animal-bacterial association. *Appl. Environ. Microbiol.* **70,** 1434–1441
44. Aeckersberg, F., Lupp, C., Feliciano, B., and Ruby, E. G. (2001) Vibrio fischeri outer membrane protein OmpU plays a role in normal symbiotic colonization. *J. Bacteriol.* **183,** 6590–6597
45. Duperthuy, M., Schmitt, P., Garzon, E., Caro, A., Rosa, R. D., Le Roux, F., Lautredou-Audouy, N., Got, P., Romestand, B., de Lorgeril, J., Kieffer-Jaquinod, S., Bachere, E., and Destoumieux-Garzon, D. (2011) Use of OmpU porins for attachment and invasion of *Crassostrea gigas* immune cells by the oyster pathogen Vibrio splendidus. *Proc. Natl. Acad. Sci.*

U.S.A. **108,** 2993–2998
46. Huang, Y., Jellies, J., Johansen, K. M., and Johansen, J. (1997) Differential glycosylation of tractin and LeechCAM, two novel Ig superfamily members, regulates neurite extension and fascicle formation. *J. Cell Biol.* **138,** 143–157
47. Jie, C., Zipser, B., Jellies, J., Johansen, K. M., and Johansen, J. (1999) Differential glycosylation and proteolytical processing of LeechCAM in central and peripheral leech neurons. *Biochim. Biophys. Acta* **1452,** 161–171
48. Johansson, M. W. (1999) Cell adhesion molecules in invertebrate immunity. *Dev. Comp. Immunol.* **23,** 303–315
49. Springer, T. A. (1994) Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. *Cell* **76,** 301–314
50. Yang, L., Johansson, J., Ridsdale, R., Willander, H., Fitzen, M., Akinbi, H. T., and Weaver, T. E. (2010) Surfactant protein B propeptide contains a saposin-like protein domain with antimicrobial activity at low pH. *J. Immunol.* **184,** 975–983
51. Chroneos, Z. C., Sever-Chroneos, Z., and Shepherd, V. L. (2010) Pulmonary surfactant: an immunological perspective. *Cell Physiol. Biochem.* **25,** 13–26
52. Vasta, G. R. (2009) Roles of galectins in infection. *Nat. Rev. Microbiol.* **7,** 424–438
53. Vasta, G. R. (2012) Galectins as pattern recognition receptors: structure, function, and evolution. *Adv. Exp. Med. Biol.* **946,** 21–36
54. Shi, X. Z., Wang, L., Xu, S., Zhang, X. W., Zhao, X. F., Vasta, G. R., and Wang, J. X. (2014) A galectin from the kuruma shrimp (*Marsupenaeus japonicus*) functions as an opsonin and promotes bacterial clearance from hemolymph. *PLoS One* **9,** e91794
55. Tasumi, S., and Vasta, G. R. (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
56. Yu, Y., Yuan, S., Yu, Y., Huang, H., Feng, K., Pan, M., Huang, S., Dong, M., Chen, S., and Xu, A. (2007) Molecular and biochemical characterization of galectin from amphioxus: primitive galectin of chordates participated in the infection processes. *Glycobiology* **17,** 774–783
57. Saftig, P., and Klumperman, J. (2009) Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat. Rev. Mol. Cell Biol.* **10,** 623–635
58. Darmoise, A., Maschmeyer, P., and Winau, F. (2010) The immunological functions of saposins. *Adv. Immunol.* **105,** 25–62
59. Zavasnik-Bergant, T., and Turk, B. (2006) Cysteine cathepsins in the immune response. *Tissue Antigens* **67,** 349–355
60. Lefebvre, C., Vandenbulcke, F., Bocquet, B., Tasiemski, A., Desmons, A., Verstraete, M., Salzet, M., and Cocquerelle, C. (2008) Cathepsin L and cystatin B gene expression discriminates immune coelomic cells in the leech *Theromyzon tessulatum. Dev. Comp. Immunol.* **32,** 795–807
61. Tryselius, Y., and Hultmark, D. (1997) Cysteine proteinase 1 (CP1), a cathepsin L-like enzyme expressed in the *Drosophila melanogaster* haemocyte cell line mbn-2. *Insect Mol. Biol.* **6,** 173–181
62. Nepal, R. M., Mampe, S., Shaffer, B., Erickson, A. H., and Bryant, P. (2006) Cathepsin L maturation and activity is impaired in macrophages harboring *M. avium* and *M. tuberculosis. Int. Immunol.* **18,** 931–939
63. Thomas, V., Samanta, S., and Fikrig, E. (2008) Anaplasma phagocytophilum increases cathepsin L activity, thereby globally influencing neutrophil function. *Infect. Immun.* **76,** 4905–4912
64. Nishikori, K., Morioka, K., Kubo, T., and Morioka, M. (2009) Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. *J. Insect Physiol.* **55,** 351–357
65. Troll, J. V., Adin, D. M., Wier, A. M., Paquette, N., Silverman, N., Goldman, W. E., Stadermann, F. J., Stabb, E. V., and McFall-Ngai, M. J. (2009) Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. *Cell Microbiol.* **11,** 1114–1127
66. Troll, J. V., Bent, E. H., Pacquette, N., Wier, A. M., Goldman, W. E., Silverman, N., and McFall-Ngai, M. J. (2010) Taming the symbiont for coexistence: a host PGRP neutralizes a bacterial symbiont toxin. *Environ. Microbiol.* **12,** 2190–2203
67. Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature* **473,** 337–342
68. Vogel, C., and Marcotte, E. M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13,** 227–232
69. Doino Lemus, J., and McFall-Ngai, M. J. (2000) Alterations in the proteome of the *Euprymna scolopes* light organ in response to symbiotic *Vibrio fischeri. Appl. Environ. Microbiol.* **66,** 4091–4097
70. Danielsen, M., Hornshoj, H., Siggers, R. H., Jensen, B. B., van Kessel, A. G., and Bendixen, E. (2007) Effects of bacterial colonization on the porcine intestinal proteome. *J. Proteome Res.* **6,** 2596–2604
71. de Lorgeril, J., Zenagui, R., Rosa, R. D., Piquemal, D., and Bachere, 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(8):** e23142
72. Neilson, K. A., Mariani, M., and Haynes, P. A. (2011) Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics* **11,** 1696–1706
73. Wang, H., Alvarez, S., and Hicks, L. M. (2012) Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two *Chlamydomonas reinhardtii* strains of interest for biofuels engineering. *J. Proteome Res.* **11,** 487–501
74. Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, Dianes JA, Sun Z, Farrah T, Bandeira N, Binz PA, Xenarios I, Eisenacher M, Mayer G, Gatto L,Campos A, Chalkley RJ, Kraus HJ, Albar JP, Martinez-Bartolomé S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H (2014). ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nature Biotechnol.* **30(3):** 223–226. PubMed PMID:24727771.
75. Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelleiro D, Perez-Riverol Y, Reisinger F, Rios D, Wang R, Hermjakob H (2013). The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41(D1):** D1063–D1069. PubMed PMID: 23203882.