Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate

Sham V. Nair, Heather Del Valle, Paul S. Gross, David P. Terwilliger and L. Courtney Smith

Physiol. Genomics 22:33-47, 2005. First published Apr 12, 2005;
doi:10.1152/physiolgenomics.00052.2005

Physiol Genomics 22: 33–47, 2005.
First published April 12, 2005; 10.1152/physiolgenomics.00052.2005.

Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate

Sham V. Nair,[1] Heather Del Valle,[1] Paul S. Gross,[2] David P. Terwilliger,[1] and L. Courtney Smith[1]

[1]*Department of Biological Sciences, George Washington University, Washington, District of Columbia; and* [2]*Department of Biochemistry, Medical University of South Carolina, Charleston, South Carolina*

Submitted 1 March 2005; accepted in final form 6 April 2005

Nair, Sham V., Heather Del Valle, Paul S. Gross, David P. Terwilliger, and L. Courtney Smith. Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin. Identification of unexpected immune diversity in an invertebrate. *Physiol Genomics* 22: 33–47, 2005. First published April 12, 2005; 10.1152/physiolgenomics.00052.2005.—The purple sea urchin, *Strongylocentrotus purpuratus*, is a member of the phylum Echinodermata, which is basal to the phylum Chordata within the deuterostome lineage of the animal kingdom. This relationship makes the analysis of the sea urchin immune system relevant to understanding the evolution of the deuterostome immune system leading to the Vertebrata. Subtractive suppression hybridization was employed to generate cDNA probes for screening high-density arrayed, conventional cDNA libraries to identify genes that were upregulated in coelomocytes responding to lipopolysaccharide. Results from 1,247 expressed sequence tags (ESTs) were used to infer that coelomocytes upregulated genes involved in RNA splicing, protein processing and targeting, secretion, endosomal activities, cell signaling, and alterations to the cytoskeletal architecture including interactions with the extracellular matrix. Of particular note was a set of transcripts represented by 60% of the ESTs analyzed, which encoded a previously uncharacterized family of closely related proteins, provisionally designated as *185/333*. These transcripts exhibited a significant level of variation in their nucleotide sequence and evidence of putative alternative splicing that could yield up to 15 translatable elements. On the basis of the striking increase in gene expression in response to lipopolysaccharide and the unexpected level of diversity of the *185/333* messages, we propose that this set of transcripts encodes a family of putative immune response proteins that may represent a major component of an immunological response to bacterial challenge.

innate; echinoderm; lipopolysaccharide; *185/333*

THE IMMUNE SYSTEM of the purple sea urchin, *Strongylocentrotus purpuratus*, lacks the hallmarks of adaptive immunity (103). Yet, as with other invertebrates studied thus far, there are several commonalities among the innate immune reactions of all animals. Current paradigms suggest that invertebrates recognize conserved molecular structures, or pathogen-associated molecular patterns (PAMPs), displayed by invading pathogens. Recognition is mediated by pattern recognition receptors (PRRs), which lack the discriminative capacity of the immunoglobulin-based receptors of the vertebrate adaptive immune system (52, 68). This paradigm is now being challenged by newly emerging data from a number of laboratories, including ours, which may necessitate a reappraisal of immune diversity in invertebrates (15, 16, 24, 63 and reviewed in Ref. 28).

Studies assessing the immune response in sea urchins have shown that the coelomic cavity is rapidly cleared of microbes, foreign cells, and other materials by the activities of phagocytic coelomocytes (reviewed in Ref. 104). Although there are several morphologically distinct classes of coelomocytes, including phagocytes, red spherule cells, colorless spherule cells, and vibratile cells (27, 54, 103), the phagocytes appear to have a central role in host defense. Phagocytes have extensive cytoskeletons and are amoeboid, macrophage-like cells that undergo significant changes in shape during chemotaxis, phagocytosis, encapsulation, clot formation, secretion, and degranulation, all of which are associated with immune responses (32, 34, 99). Analysis of expressed sequence tags (ESTs) prepared from immune-activated coelomocytes identified a simple complement system operating in sea urchins (101). The sea urchin homolog of vertebrate complement component C3 (SpC3) was the first member of the thioester-containing family of proteins identified in an invertebrate (3). It appears that SpC3 is functionally similar to its vertebrate homologs (97) in that it mediates opsonization of foreign matter and pathogens, which leads to augmented uptake by phagocytic coelomocytes (20, 96). A homolog of factor B (Bf), called SpBf, has also been identified in sea urchins (98, 113) along with two cDNAs that encode the mosaic proteins *Sp5* and *Sp5013* that share a number of domains [including a factor I membrane attack complex (FIMAC) domain] with complement regulatory proteins and members of the terminal complement pathway (70).

The use of large-scale screening technologies employing microarrays, cDNA libraries, and proteomics (118) has provided insights into transcriptional programs that are initiated in animal cells responding to pathogenic threat. These approaches have significantly accelerated the rate of gene discovery in invertebrates, including two genera of shrimp, *Penaeus* (108) and *Litopenaeus* (33); two species of oyster, *Crassostrea gigas* (35) and *C. virginica* (53); and the sea urchin, *S. purpuratus* (101). Such studies have been fruitful even for the more familiar and well-characterized invertebrates such as the fruit fly, *Drosophila melanogaster* (18, 48a), and two mosquito species, *Anopheles gambiae* (17) and *Aedes aegypti* (94). Comparative analysis of the emerging data indicates that there are general classes of genes that may be the common denominators of invertebrate immune responses. The molecules encoded by these genes include peptidoglycan recognition proteins, thioester-containing proteins, gram-negative binding proteins, multidomain scavenger receptor family, ficolins, C-

Article published online before print. See web site for date of publication (http://physiolgenomics.physiology.org).

Address for reprint requests and other correspondence: L. C. Smith, Dept. of Biological Sciences, 340 Lisner Hall, George Washington Univ., 2023 G St. NW, Washington, DC 20052 (e-mail: csmith@gwu.edu).

type lectins, and galectins. Other commonly expressed groups of proteins comprise proteases (serine proteases, cysteine proteases, metalloproteases), protease inhibitors (kazal-type serine protease inhibitor, serpin, α2-macroglobulin), antimicrobial peptides, superoxide dismutase, and metal-binding proteins. Proteolytic cascades (such as the prophenoloxidase and the clotting/coagulation systems) and certain signal transduction systems, especially those leading to the activation of rel/NF-κB (17, 35, 108, 33), are also common components of immune activation in invertebrates.

By employing EST studies or differential display, a number of genes in sea urchin coelomocytes have been shown to be induced by immune challenge, including *Sp064*, which encodes SpC3 (19), the transcription factors SpNFkB and SpRUNT (78), as well as other uncharacterized transcripts (83). The expression of a C-type lectin, *Sp056* (GenBank accession no. AY663300), is only detectible after LPS challenge and has been used as a reliable marker of coelomocyte activation (70, 113). On the other hand, there are some immune-related genes expressed in coelomocytes that are not induced by immune challenge, including *Sp152* (SpBf; Ref. 113), *Sp5* and *Sp5013* (70), and a large set of scavenger receptors with cysteine-rich (SRCR) domains (76, 77). Furthermore, the transcription factor SpGATAc has been shown to be downregulated in response to immune challenge (78).

At present, only a limited number of immune effector cascades (e.g., complement components, a set of inducible lectins, and SRCR proteins) have been identified in sea urchins, but given the effectiveness of the sea urchin to repel infection, it is expected that the sea urchin immune system will have many more as yet unidentified components and cascades. It is hypothesized that large-scale genomics-based screening approaches should reveal additional genes with central importance in sea urchin immunity. To test this hypothesis, normalized subtracted probes were used to identify expressed genes that were induced by LPS through screening two arrayed, conventional, coelomocyte cDNA libraries, one of which was constructed from bacterially activated cells. Reported here are the detailed results from an analysis of ESTs, which revealed that a number of genes were induced, including many novel ones. Matches to known sequences suggested that coelomocytes responded to gram-negative bacteria by increased expression of genes encoding proteins involved in RNA splicing, protein processing and secretion, signaling pathways, modification of the cytoskeletal architecture, and altered interactions with the extracellular matrix. Of particular note was the observation that the vast majority of the ESTs represented a set of similar transcripts that encoded a family of putative immune response proteins. It is hypothesized that these proteins may be a major immune effector system in the sea urchin based on a striking increase in gene expression in response to LPS plus significant sequence diversity due to variations in the primary nucleotide sequence and putative alternative splicing.

MATERIALS AND METHODS

Sea urchins. Animals were obtained and housed as previously described (34, 95). Immunoquiescent animals were generated by long-term housing (>6–8 mo) in a closed aquarium without significant disturbance (19, 32, 95).

Animal treatments and RNA isolation. Coelomocytes were withdrawn from sea urchins as described (32, 70), and total RNA was isolated using an RNeasy Mini kit according to the manufacturer's protocol (Qiagen, Valencia, CA). Sea urchins were injected once on *day 1* and twice on *day 2* (~9-h interval) with 0.5 μg LPS/ml coelomic fluid, according to the methods of Smith et al. (99). Coelomocytes collected for RNA isolation and cDNA subtractions were obtained before injection (immunoquiescent) and on *day 3* (LPS activated).

RT-PCR and construction of suppressive subtractive hybridization cDNA. Reverse transcription reactions with random hexamer primers were carried out on 0.5–3.0 μg of total RNA, 200 U of Superscript II RT (Life Technologies, Carlsbad, CA), and 20 U RNAsin (Promega, Madison, WI). The resultant cDNA (1 μl) was amplified by PCR using 1 U of *Taq* polymerase (Life Technologies) and 1 μM each primer (Supplemental Table S1; available at the *Physiological Genomics* web site).[1] Construction of suppressive subtractive hybridization (SSH) cDNA was carried out using the PCR-select cDNA subtraction kit (Clontech, Palo Alto, CA) according to the manufacturer's instructions. Forward cDNA subtraction employed coelomocyte RNA from an LPS-activated sea urchin (tester) subtracted from RNA prepared from the same animal before challenge (driver), while reverse cDNA subtraction used immunoquiescent RNA (tester) subtracted from LPS-activated RNA from the same sea urchin (driver). The quality of the SSH reaction was assessed using a virtual Northern blot. Samples of forward- and reverse-subtracted cDNAs, in addition to unsubtracted cDNA controls, were amplified by PCR using Clontech primers, electrophoresed, and blotted using standard methods (87). Filters were probed with inserts from specific clones (Supplemental Table S1) that were amplified by PCR and labeled with alkaline phosphatase according to the manufacturer's instructions (AlkPhos Direct; Amersham Pharmacia Biotech, Piscataway, NJ).

Riboprobe preparation. Forward- and reverse-subtracted cDNAs that served as riboprobe templates were amplified by PCR using the P1 forward primer and the nested P2R reverse primer (PCR-select cDNA subtraction kit, Clontech). The amplified cDNAs were labeled with the use of T7 RNA polymerase (Promega) and [^{32}P]rUTP (ICN, Costa Mesa, CA). The eluates from G50 Sephadex spin columns (Amersham Pharmacia Biotech) were quantified in an LS6500 scintillation counter (Beckman Instruments, Fullerton, CA). Probes were diluted to a final volume of 500 μl in hybridization solution [50% deionized formamide, 0.25 M phosphate buffer, 0.1% (wt/vol) BSA, 1 mM EDTA, 7% (wt/vol) SDS] and heated to 65°C for 4 min before addition to the filters.

Riboprobes were also generated by run-off polymerization using three *185/333* clones (*Sp0323*, *Sp0324*, *Sp0325*). Each clone was linearized with *Xho*I, and [^{32}P]rUTP (ICN) was incorporated into a riboprobe with T3, T7, or Sp6 RNA polymerase, depending on the template (see Ref. 70). Unincorporated [^{32}P]rUTP was removed with the use of a G50 Sephadex spin column (Amersham Pharmacia Biotech). Probes were heated to 60°C for 4 min before hybridization.

Macroarray filter screens and clone selection. Two arrayed cDNA libraries (5 filters per library of 92,160 clones) were used for analysis, one constructed from bacterially activated coelomocytes (14, 78, 82) and a second from nonactivated coelomocytes (3). For both libraries, coelomocytes were pooled from five sea urchins. Each library was screened with ^{32}P-labeled riboprobes derived from the SSH cDNAs and specific cloned templates that were also used for virtual Northern blot analysis (Supplemental Table S1). Briefly, macroarray filters were prehybridized for 1 h at 42°C followed by hybridization with the riboprobe at 42°C overnight in a rotating oven. Filters were washed twice for 30 min at 65°C in each of the following wash buffers: 4× SSC with 1% SDS, 2× SSC with 1% SDS, and 1× SSC with 1% SDS. Wet filters were sealed within plastic bags and exposed to film

[1] The Supplemental Material for this article (Supplemental Tables S1 and S2) is available online at **http://physiolgenomics.physiology.org/cgi/content/full/00052.2005/DC1**.

at −70°C using intensifying screens for varying periods. After autoradiography, filters were stripped of the probes by denaturation in 0.4 M NaOH for 30 min at 45°C followed by incubation in stripping buffer (0.1× SSC, 0.1% SDS, 0.2 M Tris·HCl, pH 7.5) for 30 min at 65°C. Final wash was in stripping buffer with 0.2 M EDTA for 10 min at room temperature. After stripping, filters were exposed to film to ensure that there was no residual radioactivity associated with the filters and stored at −20°C.

Clones identified during screening were manually picked, isolated, robotically arrayed into 96-well plates (as described in Ref. 33), and sequenced (Amplicon Express, Pullman, WA). Sequence similarities were identified in the nonredundant nucleic acid database and the EST database (GenBank) using basic local alignment search tool (BLAST)n and BLASTx algorithms accessed from the National Center for Biotechnology Information (NCBI) website (**http://www.ncbi.nlm.nih.gov/Blast/**). Only matches with *e*-values of $<10^{-6}$ were considered significant. The ESTs generated in this study have been submitted to GenBank (accession nos. CK828301–CK829214 and CV652690–CV652795).

Nonsynonymous-synonymous (dn/ds) ratio analysis. The 5′-ends of 293 of the best quality *185/333* ESTs were analyzed for nucleotide diversity. The region was 195 bases long and was manually aligned in BioEdit (38). Analysis of the alignment using WinClada (75) indicated that 250 of the 293 sequences were duplicates, which were removed to simplify further analysis. Two clones, *Sp0267* and *Sp0291*, that had stop codons in *element 2* were also removed, leaving 42 unique EST clones for phylogenetic analysis by maximum likelihood (PAML) (126). A neighbor-joining tree was generated in molecular evolutionary genetics analysis (MEGA) v.3.0 (56) and used for the PAML analysis.

RESULTS AND DISCUSSION

The availability of immunoquiescent sea urchins has been an important tool for the investigation of specific responses to injury and pathogenic challenge. This is a limiting factor when attempting to assess immune responses in animals collected within a few days or weeks from the wild. Contact with microbes and pathogens in a natural environment activates host defenses continuously in these animals and therefore precludes their use as baseline or downregulated controls. Stress from collection and shipping exacerbates immune activation. Hence, in this study we have used only immunoquiescent animals to generate a baseline of immunological activation. Immunoquiescence has been defined as the condition in which the abundance of certain immune-specific transcripts becomes diminished in sea urchin coelomocytes compared with animals living in their natural environment (32). Immunoquiescence has been demonstrated in *S. purpuratus* after animals have been housed in closed-circulation aquaria in the laboratory for 6–8 mo (19, 20, 32, 34). Long-term housing does not induce immunosuppression, only immunoquiescence, as immune reactivity can be reversed with injections of bacteria or LPS or from injury. Such challenges result in increases in the relative abundance of the immune-specific transcripts within individual animals (19, 20). In a modified approach to that reported previously (101), immunoquiescent sea urchins were used to obtain RNA from coelomocytes pre- and post-LPS challenge for normalized subtractions. Subtracted cDNAs generated from coelomocytes collected from the same animal were used to produce riboprobes for screening filters from arrayed coelomocyte cDNA libraries. By employing the same animal for these purposes, we avoided inherent problems associated with the genetic diversity present among out-bred sea urchins (99, 100).

Immunoquiescence and RT-PCR analysis of LPS activation. The immunoquiescent status of four randomly chosen sea urchins was assessed by RT-PCR for expression of several coelomocyte-specific genes. Two animals were found to be immunoquiescent, as evidenced by the lack of initial expression of the gene encoding a C-type lectin (*Sp056*; Fig. 1*A*) and its subsequent induction after challenge (Fig. 1*B*). This result was in agreement with previous work wherein *Sp056* proved to be a useful marker for assessing immune activation to LPS challenge (70, 113). A second gene, whose pre- and postchallenge induction is more variable among individual animals under laboratory-rearing conditions, is *Sp064,* which encodes SpC3 (19). It was found to have low expression or was undetectable before challenge (Fig. 1*A*, *animals 2* and *3*) but was inducible in all animals after LPS injection, with one individual exhibiting a large change in expression (Fig. 1*B*, *animal 2*). The expression of three additional genes, *SpNFkB* (78), *Sp152* (113) and cytoplasmic actin (*CyI*; Ref. 58), was analyzed by RT-PCR but showed no change in response to LPS (data not shown). *Animal 2*, which showed the largest detectable increase in *Sp064* expression, was chosen for further analysis.

cDNA subtraction and virtual Northern blot analysis. Forward- and reverse-subtracted cDNAs were generated by SSH from isolated RNA samples obtained from *animal 2*. Forward subtraction selectively enriches the pool of transcripts that increase upon LPS challenge and injury, whereas the reverse subtraction enriches for transcripts that diminish upon LPS challenge and injury. The efficiency of both the forward and reverse subtractions was evaluated by comparing subtracted and unsubtracted cDNAs using virtual Northern blot analysis. Reverse subtraction effectively removed *CyI* and *Sp056* transcripts from the immunoquiescent cDNA pool (Fig. 2, *lanes 3* and *11*), whereas these cDNAs were enriched in the LPS-activated or forward-subtracted cDNA preparation (Fig. 2, *lanes 4* and *12*). This indicated that successful subtraction had occurred in both directions. Both *Sp056* and *CyI* were identified in the unsubtracted cDNA from both activated and immunoquiescent coelomocytes, although greater hybridization was observed in the LPS-activated cDNA pools (Fig. 2, *lanes 2* and *10*). The change in *CyI* expression may actually reflect a real

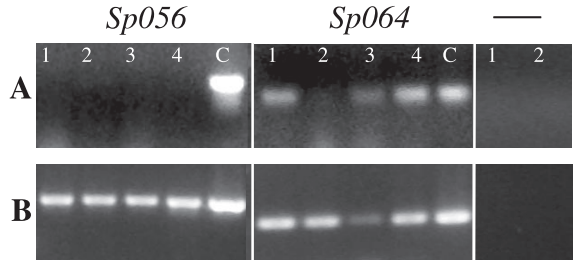

Fig. 1. Gene expression in coelomocytes from immunoquiescent sea urchins before and after LPS activation. Coelomocytes were collected from four animals (*1–4*), and gene expression was analyzed by RT-PCR before (*A*) and after (*B*) LPS challenge. Primers for *Sp056* and *Sp064* were used to amplify cDNAs (Supplementary Table S1). *Animals 2* and *3* showed increased band intensity after LPS challenge using primers for *Sp064*. None of the sea urchin samples supported amplification with *Sp056* (lectin) primers before challenge; however, all samples demonstrated expression of *Sp056* after challenge. For each set of reactions, positive controls (C) employed a cloned cDNA as the template. Negative controls (—) omitted cloned templates from the PCR reactions: *1*, *Sp056*; *2*, *Sp064*.

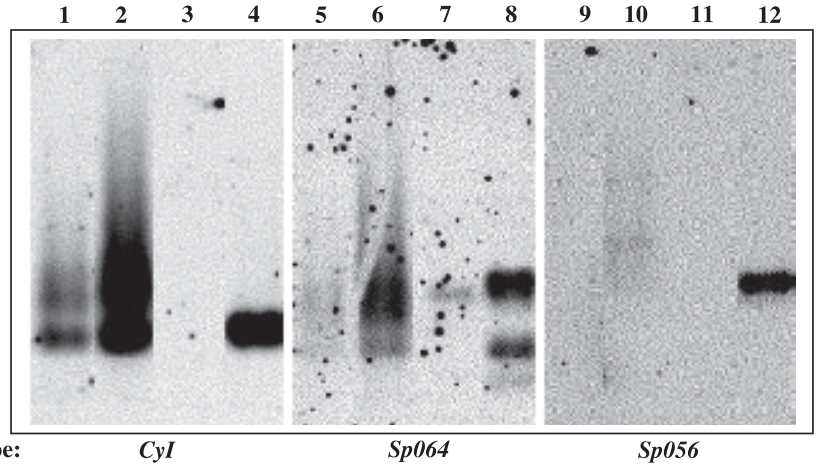

Fig. 2. Virtual Northern blot analysis of genes expressed in coelomocytes. The subtraction process was analyzed using a virtual Northern blot (see MATERIALS AND METHODS for details). Filters were probed for cytoplasmic actin (*CyI*), *Sp064*, or *Sp056* as indicated. *Lanes 1*, *5*, and *9*: unsubtracted cDNA from immunoquiescent coelomocytes. *Lanes 2*, *6*, and *10*: unsubtracted cDNA from LPS-activated coelomocytes. *Lanes 3*, *7*, and *11*: reverse-subtracted cDNA. *Lanes 4*, *8*, and *12*: forward-subtracted cDNA.

increase in actin gene expression in LPS-activated coelomocytes and would correspond with increased profilin expression (99, 100), which is involved in modifications to cytoskeletal shape (81).

Macroarray filter screens. Macroarray filters from two conventionally constructed and arrayed coelomocyte cDNA libraries were screened twice each with reverse-subtracted riboprobes and then stripped and rescreened with forward-subtracted riboprobes. Spots that were positive either for the reverse-subtracted probe or for both probes were not considered further. Filters were also screened for *CyI* (actin), *Sp056* (lectin), and *Sp064* (SpC3), and positive clones were excluded from analysis (data not shown). From ~6,000 clones that were repeatedly identified from the bacterially activated library as genes that were differentially expressed during LPS challenge and/or injury, 1,025 randomly chosen clones were sequenced initially. Approximately 73% (746) showed significant matches to hitherto uncharacterized cDNAs known as DD185 (Ref. 78, GenBank accession no. AF228877) and EST333 (Ref. 101, GenBank accession no. R62081), hereafter provisionally called *185/333*. Because there was such a high preponderance of *185/333* sequences among the clones analyzed, filters were rescreened using *185/333* riboprobes, and positives were excluded from the analysis of a subsequent 222 positive clones. Of the 1,247 clones that were analyzed in total, 71 sequences (5.7%) contained too many indeterminate base calls for successful matching, while 328 sequences (26%) were categorized as unknowns either because there was no significant similarity to any sequence (11%) or poor (*e*-value $>10^{-6}$) similarity (15%) by BLAST searches of the nonredundant database. Only 61 ESTs matched known sequences from *Strongylocentrotus* species, and these represented 23 distinct gene products. Thus this project has deposited >1,100 ESTs into the public domain. These represent novel sea urchin transcripts that are synthesized by immune-activated coelomocytes.

Sequences that could be assigned a putative identification could be further clustered by associated function (Fig. 3). The vast majority of ESTs sequenced were *185/333* transcripts, but a number of immune-related genes were also identified. In addition, a significant number of genes involved in cytoskeletal modulation (cytoskeletal architecture, motility, and cell adhesion), nuclear activities (RNA splicing and transcription), signal transduction, protease activity, and metabolism/energy generation were identified (Table 1; for a complete list see Supplemental Table S2).

185/333 encodes a putative defensive protein. Previous studies identified *185/333* as a transcript that was expressed in sea urchin coelomocytes in response to LPS (101) or to microbial challenge and physical injury (83). Evaluation of *185/333* transcript accumulation in coelomocytes by differential display and Northern blots indicated that expression was maximal at 6 h after bacterial challenge and at 12 h after injury. In agreement with these results, filter screens from both arrayed libraries indicated that the *185/333* transcripts constituted an estimate of 5,929 of the 91,920 clones (6.45%) in the bacterially activated coelomocyte library, whereas there were only 79 positive clones (0.086%) in the nonactivated coelomocyte library of the same size. On the basis of the extraordinary change in expression level in response to LPS challenge, *185/333* was investigated in detail.

Because 60% of the 1,247 ESTs in our database matched to *185/333*, the entire coding sequence of the *185/333* message could be reconstructed from these partial sequences. Optimal alignment of *185/333* ESTs could only be accomplished by inserting large gaps between matching stretches of nucleotides that delineated possible exons. This analysis indicated that the complete *185/333* translated sequence was composed of 15 elements (Fig. 4). Because not all elements were present in every *185/333* sequence, this suggested that many of the transcripts might be alternatively spliced. Nucleotide variations in the form of base substitutions were evident throughout the length of the *185/333* messages. Figure 5 illustrates these variations and shows a partial sequence that includes the 3′-end of the leader and the first element. An inspection of the nucleotide substitutions indicated that they were not randomly distributed but occurred at specific positions along the sequences. While many substitutions were conservative, a substantial number of changes resulted in amino acid changes when translated. To determine the ratio of nonsynonymous to synonymous nucleotide changes (dn/ds), 42 unique ESTs (of 293) with high-quality 5′-ends were analyzed with PAML (126). The analysis was restricted to 193 nt, which encoded the leader and *element 1*, because it maximized the number of ESTs that could be employed in the analysis and also avoided large gaps in the alignment. Results indicated that the dn/ds

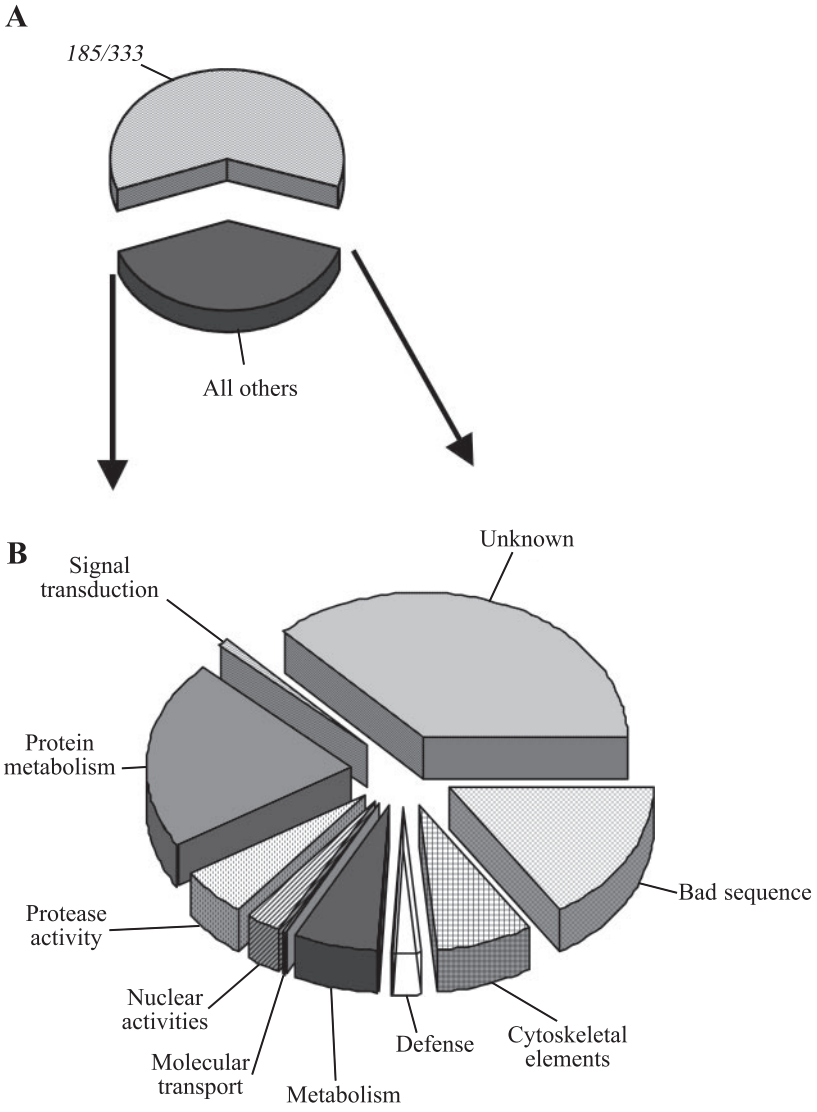

Fig. 3. A distribution of expressed sequence tags (ESTs). *A*: distribution of ESTs matching to *185/333* vs. all other ESTs. *B*: the numbers of ESTs assigned to the various functional categories were calculated as a fraction of the 1,247 ESTs minus the *185/333* matches. ESTs containing too many ambiguous base calls for satisfactory homology searching were categorized as "bad sequences." Those in the "unknown" category had significant similarities to sequences in the nonredundant databases, but had poor functional characterization. "Other" refers to sequences that did not fit into any category indicated in the pie chart.

ratio was 1.12. Any score above 1.0 implies that there are more nonsynonymous nucleotide changes than synonymous changes, and therefore, nucleotide diversity for the *185/333* genes is being maintained within the population to putatively increase fitness. However, about one-half of the sequence that was employed in the analysis encoded the leader, which appears to be less diverse than the rest of the sequence (see Fig. 4*A*). Consequently, this suggests that *element 1* was the major contributor to this result and that elevated diversity may be present in all downstream elements.

The NH_2 termini of the translated sequences contained a hydrophobic leader terminating with a Ser at position 18. The remaining protein could be divided into three regions: *1*) an NH_2-terminal glycine-rich region (Fig. 4, *A* and *C*; gly-rich, *elements 1–5*) and *2*) a histidine-rich central region (his-rich, *elements 6–14*) and a COOH-terminal region (*element 15*), which exhibited stop codons at two different positions, depending on the *185/333* sequence examined. Two of the 618 *185/333* sequences that were analyzed revealed stop codons in *element 2* (Fig. 4*A*; *Sp0269* and *Sp0289*) that would encode an extremely short protein. Nucleotide insertions/deletions were identified in the leader region resulting in two translated variants (MEVK vs. M-VK) as well as in *elements 9* and *15* (Fig. 4*A*). While no signature sequences for O-linked glycosylation were present, conserved NH_2-linked glycosylation sites (up to 12) were noted in the his-rich region, as well as one in the gly-rich region in some sequences (Fig. 4, *A* and *C*). Because no cysteines were found in any of the *185/333* coding regions, the structure of the encoded proteins would not be stabilized by disulfide bonds. In addition, a cell adhesion motif (RGD) was identified in *element 5*.

Short, repeated stretches of amino acids were identified within both the gly-rich and the his-rich regions, using Megalign (DNASTAR, Madison, WI), and by visual inspection [Fig. 4, *A* (colored brackets) and *C*, and Table 2]. These included imperfect repeats, which showed variations in their amino acid sequences (Table 2, *types 1*, *3*, *5*). Repeat *type 1* was only found in the gly-rich region and consisted of four full-length repeats and one short repeat, which was missing five residues from the NH_2-terminal end [Table 2 and Fig. 4, *A* (red brackets) and *C*]. *Repeats 2–5* were only distributed throughout the his-rich region, while the COOH-terminal

Table 1. *Partial list of ESTs*

Sp No.	Protein Match	*e*-Value	Species Match	Accession No.
		Defense-related proteins		
0014	glaectin 9	3e-24	*Rattus norvegicus*	CK828989
0434	cathepsin L	e-142	*Strongylocentrotus purpuratus*	CK828340
1107	thrombospondin	4e-53	*Homo sapiens*	CV652738
1108	cathepsin B	1e-44	*Urechis caupo*	CK828356
1186	phosphatidyleothanolamine-binding protein	4e-17	*Arabidopsis thaliana*	CV652720
1200	amssin	6e-18	*Strongylocentrotus purpuratus*	CV652694
		Nuclear activities and RNA splicing		
0576	ET putative translation product	0.67	*Mus musculus*	CV652705
0874	nuclear RNA- and DNA-Binding Protein	1e-06	*Homo sapiens*	CV652717
0896	nuclear protein NHP2	1e-40	*Homo sapiens*	CV652716
1034	LPS-induced TNF alpha	3e-12	*Mus musculus*	CV652710
1093	DNA methyl transerfase-associated protein	3e-37	*Mus musculus*	CV652701
1122	heterogeneous nuclear ribonucleoprotein R	1e-31	*Chironomus tentans*	CV652706
1157	orphan steriod hormone receptor 2	e-129	*Strongylocentrotus purpuratus*	CV652718
1176	paraspeckle protein	4e-56	*Homo sapiens*	CV652719
1184	DNA-binding (hexamer-binding) protein	1e-25	*Leishmania major*	CV652700
1188	immediate early response 5	3e-06	*Homo sapiens*	CV652708
1214	splicing factor 30	8e-22	*Homo sapiens*	CV652736
		Protein synthesis, processing, and degradation		
0057	Sec22 vesicle trafficking protein	4e-40	*Rattus norvegicus*	CV652732
0565	coated vesicle membrane protein (p24A)	2e-30	*Mus musculus*	CV652698
0694	vacuolar sorting protein 4b	3e-41	*Mus musculus*	CV652746
0744	Rab-7	e-113	*Strongylocentrotus purpuratus*	CV652728
0906	Sec61, ER transport protein	1e-49	*Strongylocentrotus purpuratus*	CV652733
1059	peptide chain release factor, subunit	e-112	*Drosophila melanogaster*	CV652747
1074	translation elongation factor	e-146	*Strongylocentrotus purpuratus*	CV652740
1094	vacuolar ATP synthase subunit G	6e-09	*Manduca sexta*	CV652745
1118	translocon associated protein gamma	9e-51	*Xenopus laevis*	CV652742
1136	ER calcistorin/Protein disulfide-isomerase	e-165	*Strongylocentrotus purpuratus*	CV652725
1141	ribosomal protein 40S	e-123	*Tripneustes gratilla*	CK829056
1197	mannose-6-phosphate receptor	4e-24	*Mus musculus*	CV652748
1209	transport protein	1e-24	*Mus musculus*	CV652734
1210	Rab5-interacting protein	8e-33	*Mus musculus*	CV652730
1228	presenilin	2e-24	*Anopheles gambiae*	CV652722
1244	translation initiation factor	4e-93	*Homo sapiens*	CV652741
1246	p24B precursor protein	4e-60	*Mus musculus*	CV652712
		Cytoskeleton and cell motility		
0029	mena neural variant	1e-36	*Mus musculus*	CV652713
0042	thymosin beta	e-112	*Strongylocentrotus purpuratus*	CV652739
0165	receptor, activated protein kinase C (RACK)	1e-67	*Biomphalaria glabrata*	CK829001
0188	integrin β-C	3e-06	*Strongylocentrotus purpuratus*	CV652709
0350	protein tyrosine phsophatase receptor type F	9e-43	*Homo sapiens*	CK829000
1027	gelsolin	2e-68	*Lumbricus terrestris*	CK828321
1044	protein tyrosine kinase 9-like protein	3e-21	*Mus musculus*	CV652726
1089	tubulin	5e-28	*Drosophila melanogaster*	CV652744
1096	Rho	6e-87	*Hemicentrotus pulcherrimus*	CV652731
1152	cofilin	7e-09	*Neurospora crassa*	CV652699
1189	avena	2e-11	*Gallus gallus*	CV652696
1195	microtubule associated protein	3e-74	*Strongylocentrotus purpuratus*	CV652714
		Cell proliferation/apoptosis		
0139	translationally controlled tumor protein	e-159	*Hemicentrotus pulcherrimus*	CK829070
1086	allograft inflammatory factor 1	3e-29	*Suberites domuncula*	CV652693
1117	polo-like kinase	e-163	*Hemicentrotus pulcherrimus*	CV652721
1158	Bax inhibitor-1	3e-31	*Paralichthys olivaceus*	CV652723

EST, expressed sequence tag.

region only had repeat *type 5*. It was interesting that the positioning of the elements and the positioning of the repeats in the gly-rich region did not correspond. Although *repeats 2–5* were not found in tandem arrays like *type 1*, a series of several different repeats were positioned in a tandem arrangement of 5-2-3-4 in *elements 14* and *15* (Fig. 4, *A* and *C*). Overall, the alignment indicated that the translated *185/333* sequences showed possible alternative splicing, nucleotide polymorphisms, internal repeats, potential NH_2-linked glycosylation sites limited mostly to the his-rich region, a variable positioning of the termination codon, and a lack of disulfide bonds.

Unexpected results were revealed from a nucleotide alignment of 290 of the *185/333* ESTs that had the least number of ambiguous bases and included the 5′-UTRs plus 190–196 nucleotides of open reading frame (ORF). This was essentially the same set that was analyzed for the dn/ds ratio above. On the basis of the sequence of the 5′-UTRs, the ESTs selected for this analysis could be allocated into 25 groups, of which 19 groups had between 2 and 101 members and 6 occurred as singletons (Table 3). Thirteen groups, each composed of five or more ESTs with identical 5′-UTRs, were analyzed using WinClada (75) to identify the linkage of an ORF sequence with a particular 5′-UTR. Results indicated that between one and eight ORFs were linked to a given 5′-UTR. If each 5′-UTR sequence was representative of expression from an individual gene, this result suggested that multiple genes with identical 5′-untranslated regions were present in the genome and probably encoded different proteins. The total number of different ORF sequences identified using this approach of analyzing 13 groups (with 5 or more sequences) was 40 (Table 3). On the other hand, when the total number of distinct ORFs was analyzed from all 290 ESTs without grouping the sequences by 5′-UTR identity, there were only 33 different ORFs. This indicated that identical ORFs were associated with different 5′-UTRs.

These data suggested that the *185/333* ESTs represented expression from a large family of closely related genes that may perhaps share sequences and are induced by immune challenge with LPS. Message heterogeneity appears to be a result of putative alternative splicing within and possibly between genes, as well as nucleotide polymorphisms within elements. The messages encode a family of similar but distinct proteins that show an unexpected level of sequence diversity, perhaps under positive selection for diversity as a result of pathogen pressure. On the basis of their expression patterns in response to LPS and the presence of a leader, these proteins may be secreted from the coelomocytes into the coelomic fluid and have an important immune function.

Defense-related proteins. A number of defense-related genes, in addition to the *185/333* family, were identified in this study. Amassin (*Sp1200*), a 75-kDa sea urchin protein that contains an olfactomedin domain, may be secreted by coelomocytes in response to injury or infection (46). The olfactomedin domain may represent an intercellular adhesion domain involved in coelomocyte-mediated clotting through amassin polymerization (reviewed in Ref. 46). Amassin may also bind to cell surface receptors, such as integrins on coelomocytes. It is noteworthy that a match to integrin-βC (*Sp0188*) was identified that has been reported previously from sea urchin embryos (13, 71). Both the βC- and βL-integrins can be detected serologically on the surface of coelomocytes (R. D. Burke, personal communication).

A galectin-like sequence (*Sp0014*) was identified that matched with both galectin-8 and galectin-9. Galectins are members of the S-type lectin subfamily that bind β-galactoside structures (36). Galectin-9 functions in the induction of apoptosis (55), acts as an eosinophil chemoattractant (67), and is possibly involved in inflammatory reactions as well as cell-cell adhesion (4). The expression of a galactose-binding lectin in coelomocytes may be involved in responses to microbial challenge, perhaps by mediating chemoattraction and cellular accumulations at points of contact with microbes displaying galactoside structures. The induction of the C-type lectin *Sp056* in coelomocytes (Fig. 1) is another example of a galactose-binding lectin that is induced by LPS (70, 113).

Two sequences (*Sp0157*, *Sp1107*) matched domains that are found in a family of proteins that include thrombospondin (TSP), properdin, semaphorin, angiogenesis inhibitor, hemicentrin, and fibulin-6 (1). Mosaic proteins with TSP domains function in cell-cell and cell-matrix interactions, cytoskeletal modifications, cell motility and aggregation, inflammatory responses, and wound healing (1). Although these activities are likely to occur in coelomocytes responding to LPS, the identification of a possible properdin homolog that might function in the ancestral sea urchin complement system is of great interest (102). In higher vertebrates, properdin functions to stabilize the formation of C3-convertase by the alternative pathway, and its expression in coelomocytes suggests that the sea urchin complement system may initiate the generation of an echinoderm C3-convertase (for review, see Refs. 98, 102). Sequence analysis of both *Sp0157* and *Sp1107* indicated that these ESTs represented two tandem TSP1 domains, and it is notable that properdin monomers also have TSP1 domains arranged in a tandem fashion (92).

A single EST (*Sp1186*) matched to the serine protease inhibitor phosphatidylethanolamine-binding protein (PEBP), which is a phospholipid-binding protein with inhibitory activity toward serine proteases (42). PEBP is highly conserved, and homologs have been identified in a number of species from the animal, plant, and archeal kingdoms. The interaction of PEBPs with Raf-1 kinase leads to the inhibition of the mitogen-activated protein (MAP) kinase signaling pathway, which suggests that these proteins may inhibit mitogenic signaling (42, 129). A number of cysteine proteases, such as cathepsin L (*Sp0434*) and cathepsin B (*Sp1108*), were identified. Cathepsins are known to function in lysosomes of mammalian macrophages and act to degrade basement membranes and the extracellular matrix of phagocytosed microbes (26, 80). In addition, cathepsins may be released from cells by secretory lysosomes to act on invading microbes in the extracellular milieu (106). Matches to these types of enzymes, as well as to arylsulfatase (EST003, EST004, EST072, EST401; Ref. 101), suggested that coelomocytes may respond to LPS challenge by synthesizing a significant number of proteases that are required for optimal functioning of the endosomal system.

RNA splicing and protein synthesis, processing, and degradation. Many sequence matches identified in this study, inferred from large-scale increases in the expression of genes that encode proteins involved in protein synthesis, processing and targeting, occur in coelomocytes after immunological challenge. There were several matches to proteins involved in the regulation of transcription, such as steroid hormone receptor (*Sp1157*), LPS-induced tumor necrosis factor-α (*Sp1034*), immediate early-response protein (*Sp1188*), and DNA methyl transferase-associated protein-1 (*Sp1093*). Other matches included proteins involved in RNA splicing, such as the ET putative translation product (*Sp0576*), nuclear RNA- and DNA-binding protein (*Sp0874*), heterogeneous nuclear ribonucleoprotein R (*Sp1122*), splicing factor 30 (*Sp1214*), paraspeckle protein (*Sp1176*), and a conserved nonhistone nucleic acid-binding protein (*Sp0896*). Increased expression of proteins associated with RNA splicing in coelomocytes suggested an overall increase in the expression of genes that

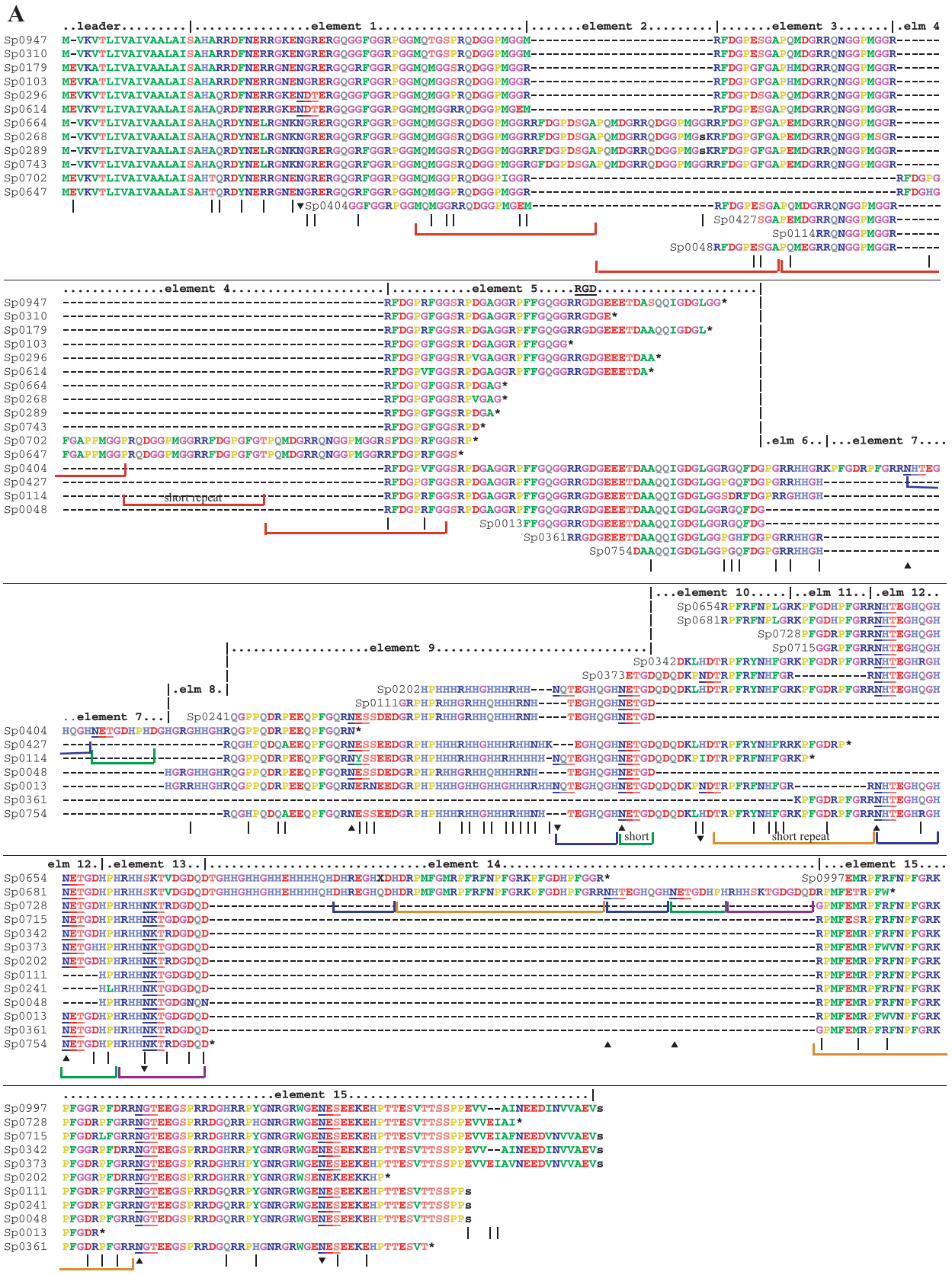

Fig. 4.

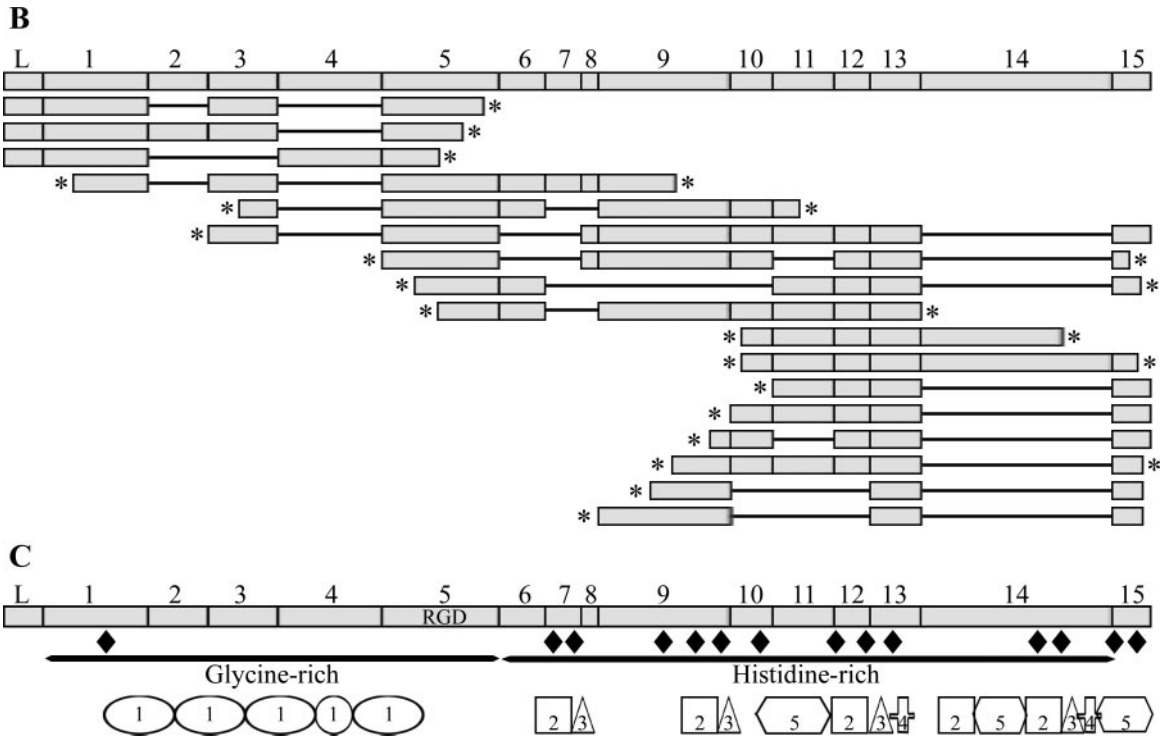

Fig. 4—*Continued*. Full-length protein alignment deduced from *185/333* EST sequences. *A*: ESTs matching to *185/333* (618 sequences) were aligned manually using BioEdit (38). Poor sequences and some duplicates were excluded, resulting in an alignment that was optimal for illustration purposes. No ESTs were full-length, and all are denoted with the clone number at the 5′-end. *Incomplete 3′-ends. Color coding of the amino acids is based on the default designation in BioEdit. Gaps are indicated by dashed lines. Each element is numbered above the alignment. The glycine-rich region includes *elements 1–5*, and the histidine-rich region includes *elements 6–14*. Internal and terminal stop codons are indicated with a lowercase "s." Vertical lines below the alignment indicate positions of variability. An RDG sequence in *element 5* is indicated above the alignment. NH_2-linked glycosylation sites are denoted within the alignment by underlining the conserved sites. They are also marked below the alignment (▼) to indicate that only some of the sequences have an Asn residue that could serve as an oligosaccharide acceptor; ▲ indicates that all sequences in the alignment have a conserved NH_2-linked glycosylation site. Amino acid repeats are identified below the alignment with colored brackets and are listed in Table 2. Incomplete repeats are labeled as "short" on the alignment. X, unknown amino acid. For reference, all elements are shown at *top* with the associated numbers. L indicates the leader. *B*: the alignment shown in *A* is illustrated as squares or rectangles sized to the relative length of each element. Lines connecting the squares or rectangles indicate where gaps were introduced to optimize the alignment. As in *A*, no EST was long enough to include the entire open reading frame. *Missing sequence at one or both ends. *C*: details of the structures and modifications of the deduced proteins are illustrated. ◆, Positions of conserved sites for NH_2-linked glycosylation. An RGD site is shown within *element 5*. The glycine-rich and histidine-rich regions are indicated with arrows. Repeats are shown as symbols and are numbered as in *A* and Table 2.

require postsynthetic RNA processing. Alternative splicing has been characterized in the expression of *Sp152*, which encodes the complement homolog SpBf, where one or two of the five short consensus repeats may be spliced out (113). Two transcripts encoding mosaic proteins with FIMAC domains may also undergo alternative splicing (70). It is possible that splicing may also be involved in *185/333* message production.

Matches to proteins involved in protein synthesis included ribosomal proteins (*Sp1141*, others), translation initiation factor (*Sp1244*), translation elongation factor (*Sp1074*), and peptide chain release factor subunit 1 (*Sp1059*). In addition, matches were identified to Sec61 (*Sp0906*) and the γ-subunit of the signal sequence receptor or the translocon-associated protein (TRAP; *Sp1118*). TRAP-γ functions in a complex that binds the signal sequence of proteins that are subsequently produced by ribosomes associated with the endoplasmic reticulum (ER) and fed through the translocation apparatus into the ER lumen (39). Sec61 is a component of the translocation channel through which nascent peptides pass (117). These matches suggested that coelomocytes responding to immune challenge may upregulate protein production, including translation on both free ribosomes and those associated with the ER.

Two protein chaperones were identified: presenilin (*Sp1228*) and the ER-localized calcistorin/protein disulfide isomerase (ERcalcistorin/PDI; *Sp1136*). Presenilin is a member of a family of aspartyl proteases that are multipass transmembrane proteins found in rough ER that are involved in processing membrane proteins (e.g., signal peptide removal, amino acid proteolysis of loops extending out from the membrane, and cleaving transmembrane regions) (125, 123). ERcalcistorin/PDI is a putative calcium storage protein in sea urchin eggs that is also found in the ER (65). Its capacity to bind Ca^{2+}, with consequent changes in the conformation of the protein, may be important for its role as a Ca^{2+}-storage protein as well as a chaperone for proper folding of newly synthesized proteins in the ER (57). The identification of protein chaperones in coelomocytes was further evidence of the upregulation of the general cellular machinery for producing and processing proteins in response to immune challenge.

There were a number of matches to proteins involved in protein sorting and trafficking of transport vesicles, including Sec22 (*Sp0057*), p24A (*Sp0565*), p24B (*Sp1246*), transport protein (*Sp1209*), vacuolar protein sorting-4b (VPS4b; *Sp0694*), and Rab-7 (*Sp0744*). The Rab family of proteins

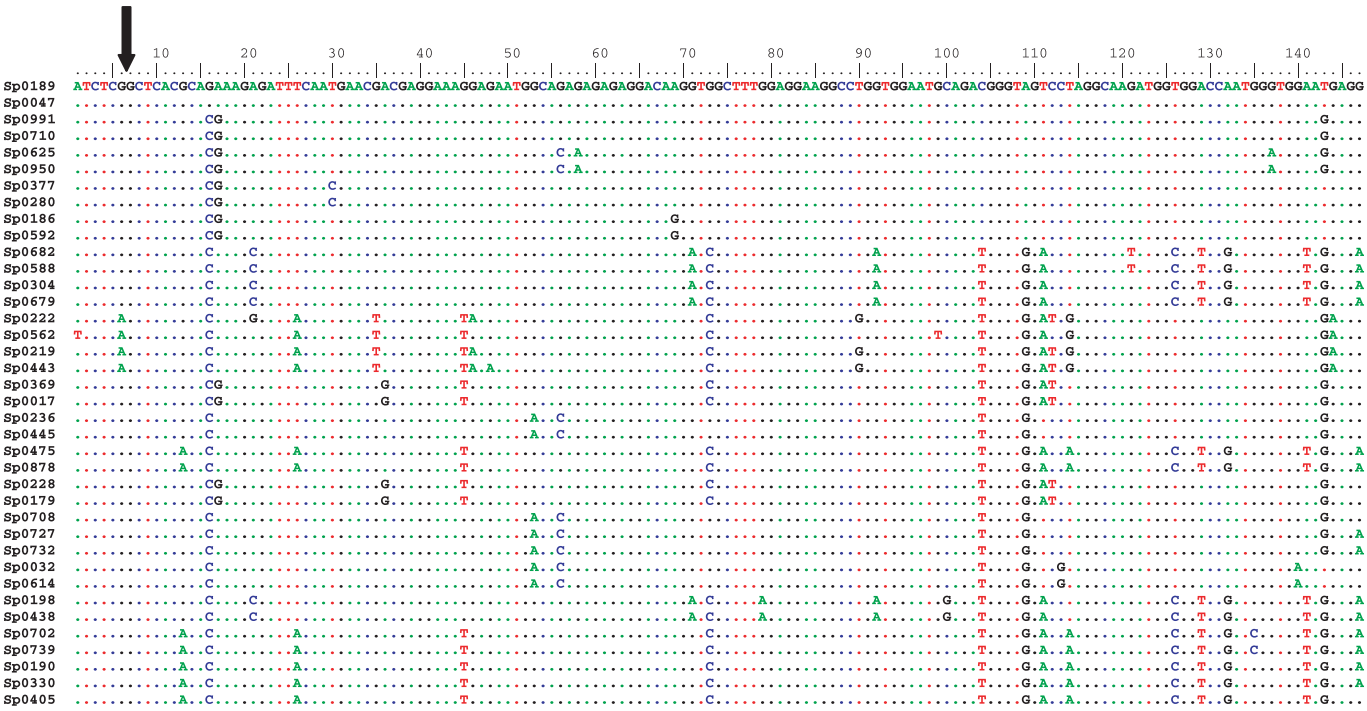

Fig. 5. Nucleotide alignment of *185/333* ESTs shows significant sequence variation. The highest quality and longest *185/333* ESTs (n = 290) were used in an alignment of sequences that included the 5′-UTR and the open reading frame encoding the leader and first element of the protein. The region shown includes the 3′-end of the leader (arrow) and the entire first element (see Fig. 4 for element locations). Duplicate sequences were reduced to two each for illustration purposes to show variations in nucleotide sequence. Periods (.) indicate conserved nucleotides, and sequence changes (for ESTs Sp0047 to Sp0405, relative to Sp 0189) are shown as nucleotides. For example, all ESTs contain an adenine at position 1, except for Sp0562, which has a thymidine at that position.

consists of small GTPases that function as major regulators of intracellular vesicular traffic (121, 130). Sec22, a v-*N*-ethylmaleimide-sensitive factor attachment protein receptor (vSNARE) protein and a member of the p24 protein family, regulates the direction and targeting of transport vesicles traveling between the ER and the Golgi apparatus (11, 22, 41, 51, 62, 64, 90, 91). VPS4b and Rab-7, on the other hand, direct vesicular traffic between the cell surface and the endosomal system (12, 10, 89, 79, 112). Several matches were identified to proteins involved in the endosomal system, including Rab5-interacting protein (*Sp1210*), which functions in the fusion and maturation of endosomal elements (47, 86, 124); the mannose-

Table 2. *185/333 Repeats*

Type 1 (red brackets)

MQMGGSRQDGGPMGGRRFDGPDSGA
PQMDGRRQDGGPMGGRRFDGPGFGA
PEMDGRRQNGGPMGGRFGAPPMGGP
RQDGGPMGGRRFDGPGFGT
PQMDGRRQNGGPMGGRRFDGPVFGG

Type 2 (blue brackets)

NHTEGHQGH
NQTEGHQGH
NHTEGHQGH
DHREGHQDH
NHTEGHQGH

Type 3 (green brackets)

NETGDHPH
NETGD
NETGDHPH
NETGDHPH

Type 4 (purple brackets)

RHHSKTVDGDQD
RHHSKTGDGDQD

Type 5 (orange brackets)

DRPMFGMRPFRFNPFGRKPFGDHPFGRR
DRPMFETRPFRFNPFGRKPFGGRPFDRR
TRPFRFNHFGR-PFGDHPFGRR

See brackets in Fig. 4*A* and numbered symbols in Fig. 4*C*.

Table 3. *Comparison of 185/333 coding* sequences†*

Group No.‡	No. of Clones	Length, nt	No. of Different Sequences§
2	38	193	8
4	101	190	8
6	9	196	1
8	5	196	2
9	23	196	6
11	9	196	4
13	12	193	1
14	16	196	1
15	5	196	2
16	6	196	1
17	11	196	2
18	25	196	3
19	7	196	1
13 groups	**267 clones**		**40 different sequences**

*The leader and the first element (see Fig. 4). †Data excluded from this analysis were 6 groups with 4 or fewer members and 6 unique sequences. ‡Groups were defined based on clustering of the 5′-UTR sequences using PAUP (109). The 5′-UTR was not used in the analysis of the coding region. §The no. of different sequences in each group was calculated in WinClada (75) after subtracting the number of sequences that were different based on ambiguous bases.

6-phosphate receptor (*Sp1197*), which targets proteins for delivery to the endosomal system (31); and a subunit of the vacuolar H^+-ATPase (*Sp1094*), which acidifies the contents of late endosomes and lysosomes (40, 60, 74, 107). Overall, the matches to proteins involved in regulating vesicle transport, trafficking, and targeting and in endosomal function suggested that coelomocytes responding to LPS increased the production of proteins that were processed through the ER, the Golgi apparatus and the endosomal system, thereby requiring increased activities in vesicular traffic.

Matches that suggested significant protein production and targeting to the endosomal system were noteworthy in light of expression of cathepsin B (*Sp1108*), cathepsin L (*Sp0434* and others; EST052; Ref. 101), cathepsin S (EST118; Ref. 101), and arylsulfatase (EST003, EST004, EST072, EST401; Ref. 101), which are proteases that are typically directed toward and function in the endosomal system. Furthermore, the most prevalent ESTs identified in this study, *185/333*, all have leader regions suggesting that they may also be processed through the ER/Golgi system and targeted to either the endosomal system or the plasma membrane. Increased secretory activity in response to LPS has been inferred from immunolocalization studies showing that two subsets of phagocytes synthesize and store the complement homolog, SpC3, in small cytoplasmic vesicles (34). After LPS challenge, significant increases in SpC3 occur in the coelomic fluid within 15 min postinjection (19, 102), indicating significant secretory activity. Overall, these data indicate that secretion and endosomal activity are significantly upregulated in responses to LPS.

Cell motility and the cytoskeleton. The majority of coelomocytes are macrophage-like, amoeboid, phagocytic cells that appear to be actively engaged in secretion and have extensive and malleable cytoskeletons (43, 44). Many of the coelomocyte ESTs identified in this study matched to structural components of the cytoskeleton or to proteins involved in modifying cytoskeletal shape with consequent changes in cell behavior. These included actin (*Sp1170* and others) and actin-binding proteins such as profilin (99, 100), cofilin (*Sp1152*), thymosin-β (*Sp0042*), and gelsolin (*Sp1027*). Protein tyrosine kinase-9 (*Sp1044*), which has a conserved ATP-binding region and a cofilin-like domain, may function in depolymerizing actin microfilaments (9). Matches to the mouse mena protein (*Sp0029*) and the chicken avena protein (*Sp1189*) were also identified. These proteins are both proline-rich cytoplasmic proteins, and mena is known to bind profilin, perhaps promoting actin polymerization (30). The presence of mena and avena homologs in coelomocytes implies active changes in cytoskeletal shape and cellular motility, which has been suggested previously based on increased expression of sea urchin profilin (99, 100).

Matches to cell surface receptors included proteins that modulate the cytoskeleton and regulate links to the extracellular matrix. For example, liprin-α2 (*Sp0350*) is believed to be involved in cell-cell and cell-matrix interactions, based on the similarity of its extracellular domain to adhesion proteins (93), while integrins (integrin-βC, *Sp0188*; Refs. 13, 71) are localized to focal adhesions, where they are known to interact with the extracellular matrix and are important in cell migration (49, 50). Furthermore, integrins interact with β-catenin and G-cadherin, which have been identified in sea urchin embryos (69). Together, expression of this set of genes, which encodes cell surface and cytoskeletal proteins, suggested interactions between the cell and the extracellular environment, perhaps inducing active modulation of the cell shape and behavior in coelomocytes.

Signal transduction, which links cell surface events to intracellular pathways, underpins the responsiveness of coelomocytes to immune challenge. Several components that function in a range of signaling systems were identified in this study, including the receptor for activated protein kinase C (RACK1; *Sp0165*), which is a member of the G protein (β-subunit) superfamily of proteins and functions at signaling nodes (128). RACK1-mediated signaling is known to affect cytoskeletal rearrangements that manifest in cell adhesion, formation of lamellipodia, and cell migration (23). In coelomocytes, the RACK1 homolog may be a key regulator of chemotactic responses to LPS. Rho GTPase (*Sp1096*), a member of the Ras family of proteins, interacts with a multitude of proteins and is another key signal transducer involved in modulating the cytoskeleton (21, 37, 66, 72, 110, 111) in addition to mediating changes in transcriptional regulation, protein translation, proliferation, motility, apoptosis, and membrane trafficking (21, 116). The deactivation of Rho proteins is mediated by GTPase-activating proteins (GAPs) and GDP dissociation inhibitors (84), which has been identified previously (EST229; Ref. 101) The expression of Rho GTPase in activated coelomocytes may be involved in coordinating a variety of signaling pathways and cellular responses to immune challenge.

Cell proliferation. A finely tuned balance exists between cell proliferation and apoptosis in the vertebrate immune response, and this may hold true for sea urchins as well. In terms of proliferation, *Sp1117* matched to polo-like kinase, a serine/threonine protein kinase known to participate in a number of activities associated with mitosis, including mitotic spindle organization and centrosome maturation. Furthermore, polo-like kinase is known to interact with β-tubulin (*Sp1089*), microtubule-associated protein (*Sp1195*), and microtubule-stabilizing proteins including the translationally controlled tumor protein (*Sp0139*; Refs. 61, 127). In addition, *Sp1188* matched to immediate early response protein-5, a nuclear protein that regulates cellular responses to mitogenic signals (122), while *Sp1086* matched to allograft inflammatory factor-1, a cytoplasmic Ca^{2+}-binding protein expressed in response to injury (7) whose overexpression has been linked to increased cell proliferation (5, 6). *Sp1184* matched to a family of DNA-binding proteins with known helicase activity that includes a DNA-binding protein characterized from *Leishmania major* (120), a DEAD/box transcription factor (85) identified in several species of tunicates (29), and the cellular nucleic acid-binding protein, a conserved protein found in a variety of vertebrates (119). The presence and increased expression of helicases could be interpreted as an indicator of DNA replication leading to mitosis. On the other hand, proliferative responses may actually be due to decreased apoptosis. A match to Bax inhibitor-1 (*Sp1158*), which is conserved in animals, fungi and plants, suggested that the apoptotic pathway in coelomocytes may be inhibited (48). Overall, the increased expression of genes encoding proteins involved in proliferative and apoptotic pathways is in general agreement with prior studies indicating that numbers of coelomocytes in the coelomic fluid increase in response to injections of LPS (19).

In summary, the change from an immunoquiescent to an immunologically activated state significantly alters the transcriptional program in coelomocytes and invokes a wide range of cellular activities. Sea urchins appear to respond to immunological challenges by active cellular remodeling of their circulating coelomocytes, with commensurate transcriptional upregulation of genes functioning in RNA processing, protein synthesis, protein processing, and trafficking. Among the genes identified in this study that appear to be involved in trafficking, vesicle fusion, and protein sorting, many of the encoded proteins may be targeted to or function within the endosomes and lysosomes, indicating increased activity of these cellular compartments.

The most striking finding among the genes inferred to be upregulated in response to LPS was the identification of a set of closely related transcripts, designated *185/333*, which were the most abundant single transcript species in immune-activated coelomocytes. The *185/333* transcripts exhibited high levels of sequence variability and putative alternative splicing of the mRNAs, resulting in encoded proteins with a wide range of size and sequence diversity. On the basis of the observed diversity and the striking increase in expression, as illustrated by the vast difference in the number of *185/333* clones in the two cDNA libraries and from Northern blots (78), a putative function of the encoded proteins is likely to include some aspect of immunity. The implication from the dn/ds ratio for the sequence at the 5′-end of the messages suggests that the genes may be under positive selection to diversify, and we speculate that it may be pathogen pressure that underlies this observation. It is not currently known how many *185/333* genes are present in the genome of an individual purple sea urchin. The data presented here give the impression that there may be hundreds; however, this may have been a result of the library structure, which used pooled coelomocytes from five sea urchins (see MATERIALS AND METHODS). Preliminary and ongoing efforts to quantitate the gene numbers have been based on a variety of approaches. Southern blots indicate that there are not hundreds of genes in a single genome; however, if the *185/*333 genes are linked, interpretation of high-molecular-weight bands is difficult (unpublished observation). Comparisons among full-length cDNA sequences from four individual sea urchins responding to LPS have indicated that each animal generates between 11 and 30 different messages based on a combination of both element usage patterns and sequence diversity within elements (unpublished observation). Searches of the trace archives of genome sequences maintained at Caltech (**http://issola.caltech.edu/~t/su-traces/blast.cgi**) and a direct analysis of the sea urchin genome (**http://www.hgsc.bcm.tmc.edu/projects/seaurchin/**) have indicated that there are 17 different sequences encoding part of the 5′-UTR plus the leader (unpublished observation). Together, these analyses have indicated that there may be as few as 8 and perhaps as many as 15 loci, and that the number of loci may vary among individual animals. Although our current estimates of gene numbers may be low, as implied from "mixing" between the 5′-UTR and the coding region sequences (Table 3), our preliminary results indicate that there are too few genes to explain the observed nucleotide variability in the ESTs. This suggests that there may be mechanisms for generating sequence diversity in the *185/333* transcripts that have not been previously characterized. On the basis of the similarities among the *185/333* cDNA sequences (and presumably among the genes), speculations on the mechanisms to generate diversity might include gene conversion, gene duplication, and RNA or DNA editing in addition to possible alternative splicing.

The unexpected level of diversity among the *185/333* sequences demonstrates a significantly greater level of complexity in the sea urchin immune system than had been imagined previously. Yet, there are other examples of nucleotide polymorphisms and significant sequence diversity in immune genes from other invertebrates. The sperm receptor located on egg surfaces of the tunicate *Halocynthia roretzi*, which mediates self sterility in these hermaphroditic animals, is highly variable and is speculated to function in allorecognition (88). Fibrinogen-related peptides (FREPs), which show significant diversity, are expressed by freshwater snail hemocytes responding specifically to the presence of parasitizing trematode worms in their hemocoel (45, 59, 131, 132). Current data suggest that FREP sequence diversity may be a result of both alternative splicing and gene conversion (63). In shrimp, the predominant transcript (6.5–21%) in hemocyte cDNA libraries was from a single family of antimicrobial peptides, the penaeidins, which exhibit significant sequence diversity, giving rise to multiple classes and isoforms within each class (25, 33). Different classes of penaeidins have differing effectiveness against specific microbial and fungal pathogens (24). Hyperdiversity has also been identified in an immunoglobulin-type variable region of an innate immune receptor in the cephalochordate *Branchiostoma floridae*, which may have chitin-binding activity (15, 16). The diversity shown in the innate immune responses of the sea urchin, snail, shrimp, and *Branchiostoma* responding to bacterial, parasitic, fungal, and viral challenges suggests that these animals, and perhaps most animals, may have hitherto unrecognized mechanisms to diversify their responses to foreignness. These mechanisms may either result in broad protection against pathogens or in directed expression of specific peptides to combat specific infecting microbes. The analysis of the sea urchin system promises to uncover mechanisms that generate diversity in the immune response, the results of which will contribute to a paradigm shift in our understanding of invertebrate immunity, as suggested by Flajnik and Du Pasquier (28). Identification of novel mechanisms for generating immune diversity in invertebrates, which has implications for innate immune capabilities in all animals, may result in a better understanding of innate immunity in higher vertebrates.

ACKNOWLEDGMENTS

We acknowledge and thank Dr. Eleanor Shepard, who assisted with the subtractions, and Cynthia Tallent, Sohum Mehta, and students in Cell Biology at George Washington University, who assisted with BLAST searches. Dr. Lakshmi Vasudevan helped to modify the Microsoft Access database program to serve our purposes. Drs. Gary Litman, David Raftos, Virginia Brockton, and Katherine Buckley provided constructive improvements to the manuscript.

Present address of S. V. Nair: Dept. of Biological Sciences, Macquarie Univ., North Ryde, NSW, 2109, Australia.

GRANTS

This research was supported by funding from the National Science Foundation (MCB-0077970 to L. C. Smith and EPS-0083102 to P. S. Gross).

REFERENCES

1. **Adams JC and Lawler J.** The thrombospondins. *Int J Biochem Cell Biol* 36: 961–968, 2004.
2. **Aicher B, Lerch MM, Muller T, Schilling J, and Ullrich A.** Cellular redistribution of protein tyrosine phosphatases LAR and PTPsigma by inducible proteolytic processing. *J Cell Biol* 138: 681–696, 1997.
3. **Al-Sharif WZ, Sunyer JO, Lambris JD, and Smith LC.** Sea urchin coelomocytes specifically express a homologue of the complement component C3. *J Immunol* 160: 2983–2997, 1998.
4. **Asakura H, Kashio Y, Nakamura K, Seki M, Dai S, Shirato Y, Abedin MJ, Yoshida N, Nishi N, Imaizumi T, Saita N, Toyama Y, Takashima H, Nakamura T, Ohkawa M, and Hirashima M.** Selective eosinophil adhesion to fibroblast via IFN-γ-induced galectin-9. *J Immunol* 169: 5912–5918, 2002.
5. **Autieri MV, Carbone C, and Mu A.** Expression of allograft inflammatory factor-1 is a marker of activated human vascular smooth muscle cells and arterial injury. *Arterioscler Thromb Vasc Biol* 20: 1737–1744, 2000.
6. **Autieri MV and Carbone CM.** Overexpression of allograft inflammatory factor-1 promotes proliferation of vascular smooth muscle cells by cell cycle deregulation. *Arterioscler Thromb Vasc Biol* 21: 1421–1426, 2001.
7. **Autieri MV, Kelemen SE, and Wendt KW.** AIF-1 is an actin-polymerizing and Rac1-activating protein that promotes vascular smooth muscle cell migration. *Circ Res* 92: 1107–1114, 2003.
8. **Banfield MJ, Barker JJ, Perry ACF, and Brady RL.** Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a role in membrane signal transduction. *Structure* 6: 1245–1254, 1998.
9. **Beeler JF, Patel BK, Chedid M, and LaRochelle WJ.** Cloning and characterization of the mouse homolog of the human A6 gene. *Gene* 193: 31–37, 1997.
10. **Beyer A, Scheuring S, Muller S, Mincheva A, Lichter P, and Kohrer K.** Comparative sequence and expression analyses of four mammalian VPS4 genes. *Gene* 305: 47–59, 2003.
11. **Blum R, Feick P, Puype M, Vandekerckhove J, Klengel R, Nastainczyk W, and Schulz I.** Tmp21 and p24A, two type I proteins enriched in pancreatic microsomal membranes, are members of a protein family involved in vesicular trafficking. *J Biol Chem* 271: 17183–17189, 1996.
12. **Buczynski G, Bush J, Zhang L, Rodriguez-Paris J, and Cardellit J.** Evidence for a recycling role for Rab7 in regulating a late step in endocytosis and in retention of lysosomal enzymes in *Dictyostelium discoideum*. *Mol Biol Cell* 8: 1343–1360, 1997.
13. **Burke RD, Murray G, Rise M, and Wang D.** Integrins on eggs: the βC subunit is essential for formation of the cortical actin cytoskeleton in sea urchin eggs. *Dev Biol* 265: 53–60, 2004.
14. **Cameron RA, Mahairas G, Rast JP, Martinez P, Biondi TR, Swartzell S, Wallace JC, Poustka AJ, Livingston BT, Wray GA, Ettensohn CA, Lehrach H, Britten RJ, Davidson EH, and Hood L.** A sea urchin genome project: Sequence scan, virtual map, and additional resources. *Proc Natl Acad Sci USA* 97: 9514–9518, 2000.
15. **Cannon JP, Haire RN, and Litman GW.** Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. *Nat Immun* 3: 1200–1207, 2002.
16. **Cannon JP, Haire RN, Schnitker N, Mueller MG, and Litman GW.** Individual protochordates have unique immune-type receptor repertoires. *Curr Biol* 14: R465–R466, 2004.
17. **Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, Meister S, Michel K, Moita LF, Müller H, Osta MA, Paskewitz SM, Reichhart J, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, Mering C, Xu J, Zheng L, Bork P, and Kafatos FC.** Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298: 159–165, 2002.
18. **Clay FJ, McEwen SJ, Bertoncello I, Wilks AF, and Dunn AR.** Identification and cloning of a protein kinase-encoding mouse gene, Plk, related to the polo gene of *Drosophila*. *Proc Natl Acad Sci USA* 90: 4882–4886, 1993.
19. **Clow LA, Gross PS, Shih SC, and Smith LC.** Expression of SpC3, the sea urchin complement component, in response to lipopolysaccharide. *Immunogenetics* 51: 1021–1033, 2000.
20. **Clow LA, Raftos DA, Gross PS, and Smith LC.** The sea urchin complement homologue, SpC3, functions as an opsonin. *J Exp Biol* 207: 2147–2155, 2004.
21. **Coleman ML, Marshall CJ, and Olson MF.** Ras and Rho GTPases in G1-phase cell-cycle regulation. *Nat Rev Mol Cell Biol* 5: 355–366, 2004.
22. **Cosson P and Letourneur F.** Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science* 263: 1629–1631, 1994.
23. **Cox E, Bennin D, Doan A, O'Toole T, and Huttenlocher A.** RACK1 regulates integrin-mediated adhesion, protrusion, and chemotactic cell migration via its src binding site. *Mol Biol Cell* 14: 658–669, 2003.
24. **Cuthbertson BJ, Bullesbach EE, Fievet J, Bachere E, and Gross PS.** A new class (penaeidin class 4) of antimicrobial peptides from the Atlantic white shrimp (*Litopenaeus setiferus*) exhibits target specificity and an independent proline-rich-domain function. *Biochem J* 381: 79–86, 2004.
25. **Cuthbertson BJ, Shepard EF, Chapman RW, and Gross PS.** Diversity of the penaeidin antimicrobial peptides in two shrimp species. *Immunogenetics* 54: 442–445, 2002.
26. **Delaisse JM, Ledent PA, and Vaes G.** Collagenolytic cysteine proteinases of bone tissue: cathepsin-B, (pro)cathepsin-L and cathepsin-L-like 70-kDa proteinase. *Biochem J* 279d: 167, 1991.
27. **Edds KT.** Effects of cytochalasin and colcemid on cortical flow in coelomocytes. *Cell Motil Cytoskeleton* 26: 262–273, 1993.
28. **Flajnik MF and Du Pasquier L.** Evolution of innate and adaptive immunity: can we draw a line? *Trends Immunol* 25: 640–644, 2004.
29. **Fujimura M and Takamura K.** Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific expression in the germ cells and its mRNA localization in the posterior-most blastomeres in early embryos. *Dev Genes Evol* 210: 64–72, 2000.
30. **Gertler FB, Niebuhr K, Reinhard M, Wehland J, and Soriano P.** Mena, a relative of VASP and *Drosophila* enabled, is implicated in the control of microfilament dynamics. *Cell* 87: 227–239, 1996.
31. **Ghosh P, Dahms NM, and Kornfeld S.** Mannose 6-phosphate receptors: new twists in the tale. *Nat Rev Mol Cell Biol* 4: 202–212, 2003.
32. **Gross PS, Al-Sharif WZ, Clow LA, and Smith LC.** Echinoderm immunity and the evolution of the complement system. *Dev Comp Immunol* 23: 429–442, 1999.
33. **Gross PS, Bartlett TC, Browdy CL, Chapman RW, and Warr GW.** Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and the Atlantic white shrimp, *L. setiferus*. *Dev Comp Immunol* 25: 565–577, 2001.
34. **Gross PS, Clow LA, and Smith LC.** SpC3, the complement homologue from the purple sea urchin, *Strongylocentrotus purpuratus*, is expressed in two subpopulations of the phagocytic coelomocytes. *Immunogenetics* 51: 1034–1044, 2000.
35. **Gueguen Y, Cadoret JP, Flament D, Barreau-Roumiguiere C, Girardot AL, Garnier J, Hoareau A, Bachere E, and Escoubas JM.** Immune gene discovery by expressed sequence tags generated from hemocytes of the bacteria-challenged oyster, *Crassostrea gigas*. *Gene* 303: 139–145, 2003.
36. **Hadari YR, Paz K, Dekel R, Mestrovic T, Accili D, and Zick Y.** Galectin-8. *J Biol Chem* 270: 3447–3453, 1995.
37. **Hall A.** Rho GTPases and the actin cytoskeleton. *Science* 279: 509–514, 1998.
38. **Hall TA.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95–98, 1999.
39. **Hartmann E, Gorlich D, Kostka S, Otto A, Kraft R, Knespel S, Burger E, Rapoport TA, and Prehn S.** A tetrameric complex of membrane proteins in the endoplasmic reticulum. *Eur J Biochem* 214: 375–381, 1993.
40. **Harvey WR.** Physiology of V-ATPases. *J Exp Biol* 172: 1–17, 1992.
41. **Hay JC, Hirling H, and Scheller RH.** Mammalian vesicle trafficking proteins of the endoplasmic reticulum and Golgi apparatus. *J Biol Chem* 271: 5671–5679, 1996.
42. **Hengst U, Albrecht H, Hess D, and Monard D.** The phosphatidylethanolamine-binding protein is the prototype of a novel family of serine protease inhibitors. *J Biol Chem* 276: 535–540, 2001.
43. **Henson JH, Kolnik SE, Fried CA, Nazarian R, McGreevy J, Schulberg KL, Detweiler M, and Trabosh VA.** Actin-based centripetal flow: phosphatase inhibition by calyculin-A alters flow pattern, actin organization, and actomyosin distribution. *Cell Motil Cytoskeleton* 56: 252–266, 2003.

44. **Henson JH, Svitkina TM, Burns AR, Hughes HE, MacPartland KJ, Nazarian R, and Borisy GG.** Two components of actin-based retrograde flow in sea urchin coelomocytes. *Mol Biol Cell* 10: 4075–4090, 1999.
45. **Hertel LA, Adema CM, and Loker ES.** Differential expression of FREP genes in two strains of *Biomphalaria glabrata* after exposure to the digenetic trematodes *Schistosoma mansoni* and *Echinostoma paraensei*. *Dev Comp Immunol* 29: 295–304, 2005.
46. **Hillier BJ and Vacquier VD.** Amassin, an olfactomedin protein, mediates the massive intercellular adhesion of sea urchin coelomocytes. *J Cell Biol* 160: 597–604, 2003.
47. **Hoffenberg S, Liu X, Nikolova L, Hall HS, Dai W, Baughn RE, Dickey BF, Barbieri MA, Aballay A, Stahl PD, and Knoll BJ.** A novel membrane-anchored Rab5 interacting protein required for homotypic endosome fusion. *J Biol Chem* 275: 24661–24669, 2000.
48. **Huckelhoven R.** BAX inhibitor-1, an ancient cell death suppressor in animals and plants with prokaryotic relatives. *Apoptosis* 9: 299–307, 2004.

48a. **Hultmark D.** Immune reactions in *Drosophila* and other insects: a model for innate immunity. *Trends Genet* 9: 178–183, 1993.

49. **Huttenlocher A, Ginsberg MH, and Horwitz AF.** Modulation of cell migration by integrin-mediated cytoskeletal linkages and ligand-binding affinity. *J Cell Biol* 134: 1551–1562, 1996.
50. **Hynes RO.** Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 69: 11–25, 1992.
51. **Jackson MR, Nilsson T, and Peterson PA.** Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO J* 9: 3153–3162, 1990.
52. **Janeway CA and Medzhitov R.** Innate immune recognition. *Annu Rev Immunol* 20: 197–216, 2002.
53. **Jenny MJ, Ringwood AH, Lacy ER, Lewitus AJ, Kempton JW, Gross PS, Warr GW, and Chapman RW.** Potential indicators of stress response identified by expressed sequence tag analysis of hemocytes and embryos from the American oyster, *Crassostrea virginica*. *Mar Biotechnol (NY)* 4: 81–93, 2002.
54. **Johnson PT.** The coelomic elements of sea urchins (*Strongylocentrotus*). I. The normal coelomocytes; their morphology and dynamics in hanging drops. *J Invertebr Pathol* 13: 42–62, 1969.
55. **Kashio Y, Nakamura K, Abedin MJ, Seki M, Nishi N, Yoshida N, Nakamura T, and Hirashima M.** Galectin-9 induces apoptosis through the calcium-calpain-caspase-1 pathway. *J Immunol* 170: 3631–3636, 2003.
56. **Kumar S, Tamura K, and Nei M.** MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5: 150–163, 2004.
57. **Lebeche D and Kaminer B.** Characterization of a calsequestrin-like protein from sea-urchin eggs. *Biochem J* 287: 741–747, 1992.
58. **Lee JJ, Shott RJ, Rose SJ, Thomas TL, Britten RJ, and Davidson EH.** Sea urchin actin gene subtypes. Gene number, linkage and evolution. *J Mol Biol* 172: 149–176, 1984.
59. **Leonard PM, Adema CM, Zhang SM, and Loker ES.** Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*. *Gene* 269: 155–165, 2001.
60. **Lepier A, Graf R, Azuma M, Merzendorfer H, Harvey WR, and Wieczorek H.** The peripheral vomplex of the tobacco hornworm V-ATPase contains a novel 13-kDa subunit G. *J Biol Chem* 271: 8502–8508, 1996.
61. **Li Q, Callaghan M, and Suprenant KA.** The 77-kDa echinoderm microtubule-associated protein (EMAP) shares epitopes with the mammalian brain MAPs, MAP-2 and Tau. *Biochem Biophys Res Commun* 250: 502–505, 1998.
62. **Liu Y and Barlowe C.** Analysis of Sec22p in endoplasmic reticulum/Golgi transport reveals cellular redundancy in SNARE protein function. *Mol Biol Cell* 13: 3314–3324, 2002.
63. **Loker ES, Adema CM, Zhang SM, and Kepler TB.** Invertebrate immune systems—not homogeneous, not simple, not well understood. *Immunol Rev* 198: 10–24, 2004.
64. **Lowe M and Kreis TE.** In vivo assembly of coatomer, the COP-I coat precursor. *J Biol Chem* 270: 31364–31371, 1995.
65. **Lucero HA, Lebeche D, and Kaminer B.** ERcalcistorin/protein disulfide isomerase (PDI). Sequence determination and expression of a cDNA clone encoding a calcium storage protein with PDI activity from endoplasmic reticulum of the sea urchin egg. *J Biol Chem* 269: 23112–23119, 1994.
66. **Mackey DJG and Hall A.** Rho GTPases. *J Biol Chem* 273: 20685–20688, 1998.
67. **Matsumoto R, Matsumoto H, Seki M, Hata M, Asano Y, Kanegasaki S, Stevens RL, and Hirashima M.** Human ecalectin, a variant of human galectin-9, is a novel eosinophil chemoattractant produced by T lymphocytes. *J Biol Chem* 273: 16976–16984, 1998.
68. **Medzhitov R and Janeway CA.** Decoding the patterns of self and nonself by the innate immune system. *Science* 296: 298–300, 2002.
69. **Miller JR and McClay DR.** Changes in the pattern of adherens junction-associated β-catenin accompany morphogenesis in the sea urchin embryo. *Dev Biol* 192: 310–322, 1997.
70. **Multerer KA and Smith LC.** Two cDNAs from the purple sea urchin, *Strongylocentrotus purpuratus*, encoding mosaic proteins with domains found in factor H, factor I, and complement components C6 and C7. *Immunogenetics* 56: 89–106, 2004.
71. **Murray G, Reed C, Marsden M, Rise M, Wang D, and Burke RD.** The α B β C integrin is expressed on the surface of the sea urchin egg and removed at fertilization. *Dev Biol* 227: 633–647, 2000.
72. **Narumiya S, Ishizaki T, and Watanabe N.** Rho effectors and reorganization of actin cytoskeleton. *FEBS Lett* 410: 68–72, 1997.
73. **Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A, and Mann M.** Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet* 20: 46–50, 1998.
74. **Nishi T and Forgac M.** The vacuolar (H^+)-ATPases—nature's most versatile proton pumps. *Nat Rev Mol Cell Biol* 3: 94–103, 2002.
75. **Nixon KC.** *WinClada Ver. 1.0000*. KC Nixon: Ithaca, NY, 1999–2002.
76. **Pancer Z.** Dynamic expression of multiple scavenger receptor cycsteine-rich genes in coelomocytes of the purple sea urchin. *Proc Natl Acad Sci USA* 97: 13156–13161, 2000.
77. **Pancer Z.** Individual-specific repertoires of immune cells SRCR receptors in the purple sea urchin (*S. purpuratus*). *Adv Exp Med Biol* 484: 31–40, 2001.
78. **Pancer Z, Rast JP, and Davidson EH.** Origins of immunity: transcription factors and homologues of effector genes of the vertebrate immune system expressed in sea urchin coelomocytes. *Immunogenetics* 49: 773–786, 1999.
79. **Pfeffer S.** Membrane domains in the secretory and endocytic pathways. *Cell* 112: 507–517, 2003.
80. **Podgorski I and Sloane BF.** Cathepsin B and its role(s) in cancer progression. *Biochem Soc Symp* 70: 263–276, 2003.
81. **Pollard TD and Borisy GG.** Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112: 453–465, 2003.
82. **Rast JP, Amore G, Calestani C, Livi CB, Ransick A, and Davidson EH.** Recovery of developmentally defined gene sets from high-density cDNA macroarrays. *Dev Biol* 228: 270–286, 2000.
83. **Rast JP, Pancer Z, and Davidson EH.** New approaches towards an understanding of deuterostome immunity. *Curr Top Microbiol Immunol* 248: 3–16, 2000.
84. **Rivero F, Illenberger D, Somesh BP, Dislich H, Adam N, and Meyer AK.** Defects in cytokinesis, actin reorganization and the contractile vacuole in cells deficient in RhoGDI. *EMBO J* 21: 4539–4549, 2002.
85. **Rocak S and Linder P.** DEAD-Box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol* 5: 232–241, 2004.
86. **Saito K, Murai J, Kajiho H, Kontani K, Kurosu H, and Katada T.** A novel binding protein composed of homophilic tetramer exhibits unique properties for the small GTPase Rab5. *J Biol Chem* 277: 3412–3418, 2002.
87. **Sambrook J, Fritsch EF, and Maniatis T.** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1989.
88. **Sawada H, Tanaka E, Ban S, Yamasaki C, Fujino J, Ooura K, Abe Y, Matsumoto K, and Yokosawa H.** Self/nonself recognition in ascidian fertilization: vitelline coat protein HrVC70 is a candidate allorecognition molecule. *Proc Natl Acad Sci USA* 101: 15615–15620, 2004.
89. **Scheuring S, Rohricht RA, Schoning-Burkhardt B, Beyer A, Muller S, Abts HF, and Kohrer K.** Mammalian cells express two VPS4 proteins both of which are involved in intracellular protein trafficking. *J Mol Biol* 312: 469–480, 2001.
90. **Schimmoller F, Singer-Kruger B, Schroder S, Kruger U, Barlowe C, and Riezman H.** The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *EMBO J* 14: 1329–1339, 1995.

91. **Schutze MP, Peterson PA, and Jackson MR.** An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. *EMBO J* 13: 1696–1705, 1994.
92. **Schwaeble WJ and Reid KBM.** Does properdin crosslink the cellular and the humoral immune response? *Immunol Today* 20: 17–21, 1999.
93. **Serra-Pages C, Kedersha NK, Fazikas L, Medley Q, Debant A, and Streuli M.** The LAR transmembrane protein tyrosine phosphatase and a coiled-coil LAR-interacting protein co-localize at focal adhesions. *EMBO J* 14: 2827–2838, 1995.
94. **Severson DW, Knudson DL, Soares MB, and Loftus BJ.** *Aedes aegypti* genomics. *Insect Biochem Mol Biol* 34: 715–721, 2004.
95. **Shah M, Brown KM, and Smith LC.** The gene encoding the sea urchin complement protein, SpC3, is expressed in embryos and can be upregulated by bacteria. *Dev Comp Immunol* 27: 529–538, 2003.
96. **Smith LC.** The complement system in sea urchins. In: *Phylogenetic Perspectives on the Vertebrate Immune Systems. Advances in Experimental Medicine and Biology*, edited by Beck M, Sugumaran M, and Cooper EL. New York: Kluwer Academic/Plenum, 2001, vol. 484, p. 363–372.
97. **Smith LC.** Thioester function is conserved in SpC3, the sea urchin homologue of the complement component C3. *Dev Comp Immunol* 26: 603–614, 2002.
98. **Smith LC, Azumi K, and Nonaka M.** Complement systems in invertebrates. The ancient alternative and lectin pathways. *Immunopharmacology* 42: 107–120, 1999.
99. **Smith LC, Britten RJ, and Davidson EH.** SpCoel1: a sea urchin profilin gene expressed specifically in coelomocytes in response to injury. *Mol Biol Cell* 3: 403–414, 1992.
100. **Smith LC, Britten RJ, and Davidson EH.** Lipopolysaccharide activates the sea urchin immune system. *Dev Comp Immunol* 19: 217–224, 1995.
101. **Smith LC, Chang L, Britten R, and Davidson EH.** Sea urchin genes expressed in activated coelomocytes are identified by expressed sequence tags. Complement homologues and other putative immune response genes suggest immune system homology within the deuterostomes. *J Immunol* 156: 593–602, 1996.
102. **Smith LC, Clow LA, and Terwilliger DP.** The ancestral complement system in sea urchins. *Immunol Rev* 180: 16–34, 2001.
103. **Smith LC and Davidson EH.** The echinoid immune system and the phylogenetic occurrence of immune mechanisms in deuterostomes. *Immunol Today* 13: 356–362, 1992.
104. **Smith LC and Davidson EH.** The echinoderm immune system: characters shared with the vertebrate immune system, and characters arising later in deuterostome phylogeny. In: *Primordial Immunity: Foundations for the Vertebrate Immune System*, edited by Beck G, Cooper EL, Habicht GS, and Marchalonls JJ. New York: New York Academy of Sciences, 1994, p. 213–226.
105. **Smith LC, Shih CS, and Dachenhausen SG.** Coelomocytes express SpBf, a homologue of factor B, the second component in the sea urchin complement system. *J Immunol* 161: 6784–6793, 1998.
106. **Stinchcombe J, Bossi G, and Griffiths GM.** Linking albinism and immunity: the secrets of secretory lysosomes. *Science* 305: 55–59, 2004.
107. **Sun-Wada GH, Wada Y, and Futai M.** Lysosome and lysosome-related organelles responsible for specialized functions in higher organisms, with special emphasis on vacuolar-type proton ATPase. *Cell Struct Funct* 28: 455–463, 2003.
108. **Supungul P, Klinbunga S, Pichyankura R, Jitrapakdee S, Hirono I, Aoki T, and Tassanakajon A.** Identification of immune-related genes in hemocytes of black tiger shrimp (*Panaeus monodon*). *Mar Biotechnol (NY)* 4: 487–494, 2002.
109. **Swofford DL.** *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4*. Sunderland, MA: Sinauer Associates, 2002.
110. **Symons M and Settleman J.** Rho family GTPases: more than simple switches. *Trends Cell Biol* 10: 415–419, 2000.
111. **Takai Y, Sasaki T, and Matozaki T.** Small GTP-binding proteins. *Physiol Rev* 81: 153–208, 2001.
112. **Tanaka H, Fujita H, Katoh H, Mori K, and Negishi M.** Vps4-A (vacuolar protein sorting 4-A) is a binding partner for a novel Rho family GTPase, Rnd2. *Biochem J* 365: 349–353, 2002.
113. **Terwilliger DP, Clow LA, Gross PS, and Smith LC.** Constitutive expression and alternative splicing of the exons encoding SCRs in Sp152, the sea urchin homologue of complement factor B. Implications on the evolution of the Bf/C2 gene family. *Immunogenetics* 56: 531–543, 2004.
114. **Utans U, Arceci RJ, Yamashita Y, and Russell ME.** Cloning and characterization of allograft inflammatory factor-1: a novel macrophage factor identified in rat cardiac allografts with chronic rejection. *J Clin Invest* 95: 2954–2962, 1995.
115. **Utans U, Quist WC, McManus BM, Wilson JE, Arceci RJ, Wallace AF, and Russell ME.** Allograft inflammatory factory-1. A cytokine-responsive macrophage molecule expressed in transplanted human hearts. *Transplantation* 61: 1387–1392, 1996.
116. **Van Aelst L and D'Souza-Schorey C.** Rho GTPases and signalling networks. *Genes Dev* 11: 2295–2322, 1997.
117. **Van den Berg B, Clemons WM, Collinson I, Modis Y, Hartmann E, Harrison SC, and Rapoport TA.** X-ray structure of a protein-conducting channel. *Nature* 427: 36–44, 2004.
118. **Vierstraete E, Verleyen P, Baggerman G, D'Hertog W, Van den Bergh G, Arckens L, De Loof A, and Schoofs L.** A proteomic approach for the analysis of instantly released wound and immune proteins in *Drosophila melanogaster* hemolymph. *Proc Natl Acad Sci USA* 101: 470–475, 2004.
119. **Warden CH, Krisans SK, Purcell-Huynh D, Leete LM, Daluiski A, Diep A, Taylor BA, and Lusis AJ.** Mouse cellular nucleic acid binding proteins: a highly conserved family identified by genetic mapping and sequencing. *Genomics* 24: 14–19, 1994.
120. **Webb J and McMaster W.** Molecular cloning and expression of a Leishmania major gene encoding a single-stranded DNA-binding protein containing nine "CCHC" zinc finger motifs. *J Biol Chem* 268: 13994–14002, 1993.
121. **Wickner W and Haas A.** Yeast homotypic vacuole fusion: a window on organelle trafficking mechanisms. *Annu Rev Biochem* 69: 247–275, 2000.
122. **Williams M, Lyu MS, Yang YL, Lin EP, Dunbrack R, Birren B, Cunningham J, and Hunter K.** Ier5, a novel member of the slow-kinetics immediate-early genes. *Genomics* 55: 327–334, 1999.
123. **Wolfe MS and Kopan R.** Intramembrane proteolysis: theme and variations. *Science* 305: 1119–1123, 2004.
124. **Woodman PG.** Biogenesis of the sorting endosome: the role of Rab5. *Traffic* 1: 695–701, 2000.
125. **Xia W and Wolfe MS.** Intramembrane proteolysis by presenilin and presenilin-like proteases. *J Cell Sci* 116: 2839–2844, 2003.
126. **Yang Z.** PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS* 13: 555–556, 1997.
127. **Yarm FR.** Plk phosphorylation regulates the microtubule-stabilizing protein TCTP. *Mol Cell Biol* 22: 6209–6221, 2002.
128. **Yarwood SJ, Steele MR, Scotland G, Houslay MD, and Bolger GB.** The RACK1 signaling scaffold protein selectively interacts with the cAMP-specific phosphodiesterase PDE4D5 isoform. *J Biol Chem* 274: 14909–14917, 1999.
129. **Yeung K, Seitz T, Li S, Janosch P, Mcferran B, Kaiser C, Fee F, Katsanakis KD, Rose D, Mischak H, Sedivy JM, and Kolch W.** Suppression of Raf-1 kinase activity and MAP kinase signalling by RKIP. *Nature* 401: 173–177, 1999.
130. **Zerial M and McBride H.** Rab proteins as membrane organizers. *Nat Rev Mol Cell Biol* 2: 107–117, 2001.
131. **Zhang S-M, Adema CM, Kepler TB, and Loker ES.** Diversification of Ig superfamily genes in an invertebrate. *Science* 305: 251–254, 2004.
132. **Zhang SM, Leonard PM, Adema CM, and Loker ES.** Parasite-responsive IgSF members in the snail *Biomphalaria glabrata*: characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain. *Immunogenetics* 53: 684–694, 2001.